



# STIC SEARCH RESULT FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact*:

Mary Hale, Information Branch Supervisor  
571-272-2507 Remsen E01 D86

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



**THIS PAGE BLANK (USPTO)**



12/498

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: PARIKIN, JEFF Examiner #: 22607 Date: 05/06/04  
 Alt Unit: 1648 Phone Number: 2-0908 Serial Number: 09/147,362  
 Mail Box and Bldg Room Location: \_\_\_\_\_ Results Format Preferred (circle) ☒ PAPER ☐ DISK ☐ E-MAIL

If more than one search is submitted, please prioritize searches in order of need. BOTH MEI

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with one appropriate serial number.*

Please search the attached claims. ~~the~~  
 We might want to simply start with SEQ ID  
 NOS.: 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, + 16.  
 If no art turns up, we might want to  
 pursue the more generic structure of claim  
 66.

Please call if you have any questions, suggestions, etc.  
 (222-0908)

exam. contacted  
 5/7/04

## STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) _____	STN _____
Searcher Phone # _____	AA Sequence (#) <u>10</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>5/7/04</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>5/10/04</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>ESP</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Fee: _____	Other _____	Other (specify): _____



Hale, Mary

---

From: Parkin, Jeffrey  
Sent: Friday, May 07, 2004 1:59 PM  
To: Hale, Mary  
Subject: RE: problem with search request for SN 09/147,362  
  
Importance: High

Mary:

Can we try the following search strategy:

- Search SEQ ID NOS.: 3, 4, 8, 9, 10, 11, 12, 13, 14, and 16 v. all relevant databases, including interference.

This should give us enough representative species to capture the generic structure.

Thanks!

JSP  
AU 1648  
REM 3D39  
2-0908

-----Original Message-----

From: Hale, Mary  
Sent: Friday, May 07, 2004 11:50 AM  
To: Parkin, Jeffrey  
Cc: Martinell, James  
Subject: re: problem with search request for SN 09/147,362

Dear Examiner Parkin,

You submitted a search request for SN 09/147,362. That request has a total of 13 sequences, which is longer than our limit of 10 sequences per search.

Your request has been cancelled. Please meet with Jim Martinell. He will review your search request and will offer suggestions for modifying the search to optimize the processing time. Once you and he have met and made appropriate changes, please resubmit the search request so we can process your request immediately. Please resend the changes or approval to search directly to me via e-mail so it can be reactivated.

If you have any questions about this, please contact Mary Hale, Supervisor, Information Branch, 2-2507

Thank you,  
Mary Hale

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 121498**

**TO: Jeffrey Parkin**  
**Location: rem/3c18**  
**Art Unit: 1648**  
**Monday, May 10, 2004**

**Case Serial Number: 09/147362**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Parkin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

THIS PAGE BLANK (CSPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 30.4538 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-10

Perfect score: 93  
Sequence: 1 LLNSWGCKRLVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	16	2	Aaw80468 Peptide d
2	93	100.0	22	2	Aaw80472 Peptide d
3	93	100.0	28	2	Aaw80473 Peptide d
4	93	100.0	33	3	Aab12231 Partial s
5	93	100.0	33	3	Aab12212 Partial s
6	93	100.0	40	2	Aaw07346 Partial s
7	93	100.0	40	2	Aaw07352 Partial s
8	91	97.8	33	3	Aab12236 Partial s
9	91	97.8	40	2	Aaw07343 Partial s
10	90	96.8	32	2	Aaw80469 Peptide d
11	89	95.7	17	2	Aaw80474 Peptide d
12	88	94.6	16	2	Aaw80467 Peptide d
13	88	94.6	22	2	Aaw80461 Peptide d
14	88	94.6	22	2	Aaw80462 Peptide d
15	87	93.5	22	2	Aaw80463 Peptide d
16	87	93.5	23	2	Aay05623 Peptide d
17	87	93.5	24	2	Aay05624 Peptide d
18	87	93.5	24	2	Aay05612 Peptide d
19	87	93.5	33	3	Aab12210 Partial s
20	87	93.5	33	3	Aab12221 Partial s
21	87	93.5	33	3	Aab12213 Partial s
22	87	93.5	41	2	Aaw07353 Partial s
23	87	93.5	42	2	Aaw07350 Partial s
24	87	93.5	116	2	Aay05555 Peptide d
25	87	93.5	117	2	Aay05548 Peptide d

26	87	93.5	715	2	Aay05625	Aay05625 HIV-1 gro
27	86	92.5	23	2	Aay05615	Aay05615 HIV-1 gro
28	86	92.5	23	2	Aay05619	Aay05619 HIV-1 gro
29	86	92.5	23	3	Aab12255	Aab12255 HIV-1 gp4
30	86	92.5	23	3	Aab12254	Aab12254 HIV-1 gp4
31	86	92.5	23	3	Aab12256	Aab12256 HIV-1 gp4
32	86	92.5	28	3	Aab12257	Aab12257 HIV-1 gp4
33	86	92.5	30	3	Aab12264	Aab12264 HIV-1 gp4
34	86	92.5	33	3	Aab12211	Aab12211 Partial s
35	86	92.5	33	3	Aab12222	Aab12222 Partial s
36	86	92.5	33	3	Aab12235	Aab12235 Partial s
37	86	92.5	33	3	Aab12214	Aab12214 Partial s
38	86	92.5	33	3	Aab12220	Aab12220 Partial s
39	86	92.5	35	3	Aab12259	Aab12259 Group O H
40	86	92.5	41	2	Aaw07351	Aaw07351 Partial s
41	86	92.5	113	2	Aay05559	Aay05559 HIV-1 gro
42	86	92.5	113	2	Aay05565	Aay05565 HIV-1 gro
43	86	92.5	115	2	Aay05557	Aay05557 HIV-1 gro
44	86	92.5	149	3	Aab12262	Aab12262 HIV group
45	86	92.5	200	3	Aay77373	Aay77373 HIV-1 gro

ALIGNMENTS

NOT PA  
RESULT 1  
AAW80468  
ID AAW80468 standard; peptide; 16 AA.  
AC AAW80468;  
XX

DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
PD 15-OCT-1998.  
XX  
PF 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
XX 24-FEB-1998; 98FR-00002212.  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human immune deficiency virus of group O.  
XX  
XX Claim 6; Page 43; 55pp; French.

CC AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences connected around short highly conserved sequences present in isolates of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 93; DB 2; Length 16;

```

Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
Db 1 LLNSWGCKGRLVCYTS 16

RESULT 2
AAW80472
ID AAW80472 standard; peptide; 22 AA.
XX
AC AAW80472;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
KW Group O human immune deficiency virus; HIV; detection; infection.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
XX
FN WO9845323-A1.
PD 15-OCT-1998.
XX
PF 06-APR-1998; 98WO-FR000691.
XX
PR 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX
DR WPI; 1998-583190/49.
XX
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
PS Claim 6; Page 44; 55pp; French.
XX
SQ AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 93; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
Db 6 LLNSWGCKGRLVCYTS 21

RESULT 3
AAW80473
ID AAW80473 standard; peptide; 28 AA.
XX
AC AAW80473;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX

Best Local Similarity 100.0%; Score 93; DB 2; Length 28;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
Db 12 LLNSWGCKGRLVCYTS 27

RESULT 4
AAB12231
ID AAB12231 standard; peptide; 33 AA.
XX
AC AAB12231;
XX
DT 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX
DE Partial sequence of HIV-1 strain MAN gp41 immunodominant region.
XX
KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; MAN.
XX
OS Human immunodeficiency virus 1.
XX
FN EP1013766-A2.
PD 28-JUN-2000.
XX
PF 29-NOV-1999; 99EP-00309491.
XX
PR 30-NOV-1998; 98US-0110292P.
PR 08-FEB-1999; 99US-0119138P.
PR 04-NOV-1999; 99US-00433428.
XX

```



PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX De Leys R, Zheng J;  
 PI WPI; 2000-402205/35.  
 DR  
 XX New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.  
 XX  
 PS Example 1; Fig 1; 52pp; English.  
 XX  
 CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus type 1 (HIV-1) strain MAN. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. MAN is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 33 AA;

Query Match 100.0%; Score 93; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRVLCVYTS 16  
 DB 14 LLNSWGCKGRVLCVYTS 29

RESULT 5  
 AAB12212  
 ID AAB12212 standard; peptide; 33 AA.  
 XX  
 AC AAB12212;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 10-NOV-2000 (first entry)  
 XX  
 DE Partial sequence of HIV-1 strain BCF13 gp41 immunodominant region.  
 XX  
 KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF13.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN EP1013766-A2.  
 XX  
 PD 28-JUN-2000.  
 XX  
 PF 29-NOV-1999; 99EP-00309491.  
 XX  
 PR 30-NOV-1998; 98US-0110292P.  
 PR 08-FEB-1999; 99US-0119138P.  
 PR 04-NOV-1999; 99US-00433428.  
 XX  
 XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX  
 PI De Leys R, Zheng J;  
 XX  
 DR WPI; 2000-402205/35.  
 XX  
 PT New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.  
 XX  
 PS Example 1; Fig 1; 52pp; English.  
 XX  
 CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain BCF13. HIV is the principle aetiological

CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
 CC envelope protein, and may be used as an antigen for the detection of  
 CC antibodies produced in response to HIV infection. BCF13 is a member of  
 CC HIV group O (outlier). The present sequence is the immunodominant region  
 CC of gp41. This sequence was used in a sequence homology alignment, which  
 CC in turn was used to derive a consensus sequence peptide: peptide 147  
 CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 33 AA;

Query Match 100.0%; Score 93; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRVLCVYTS 16  
 DB 14 LLNSWGCKGRVLCVYTS 29

RESULT 6  
 AAW07346  
 ID AAW07346 standard; peptide; 40 AA.  
 XX  
 AC AAW07346;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95FR-00002236.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Loussertajaka I, Ly T, Chaixbaudier M;  
 XX  
 DR WPI, 1996-412779/41.  
 DR N-PSDB; AAT44922.  
 XX  
 PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 34; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and WVP5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the

CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)

XX SQ Sequence 40 AA;  
 Query Match 100.0%; Score 93; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
 |||||  
 Db 19 LLNSWGCKGRLVCYTS 34

RESULT 7  
 AAW07352  
 ID AAW07352 standard; peptide; 40 AA.

XX AC AAW07352;

XX DT 16-OCT-2003 (revised)  
 XX DT 25-MAR-2003 (revised)  
 XX DT 03-JUN-1997 (first entry)

XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF13.

XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.

XX OS Human immunodeficiency virus 1.

XX PN WO9627013-A1.

XX PD 06-SEP-1996.

XX PF 26-FEB-1996; 96WO-FR000294.

XX PR 27-FEB-1995; 95FR-00002236.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX PI Simon F, Saragosti S, Lousseratajaka I, Ly T, Chaixbaudier M;  
 XX DR WPI; 1996-412779/41.

XX PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 XX PT antibodies - useful for diagnosis, screening and typing, or as  
 XX PT immunogens.

XX PS Claim 12; Page 46; 71pp; French.

XX CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and WVF5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC (BCF08 (NKO) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF13 and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)

XX SQ Sequence 40 AA;

Query Match 100.0%; Score 93; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
 |||||  
 Db 19 LLNSWGCKGRLVCYTS 34

RESULT 8  
 AAB12236  
 ID AAB12236 standard; peptide; 33 AA.

XX AC AAB12236;

XX DT 12-SEP-2003 (revised)  
 XX DT 10-NOV-2000 (first entry)

XX DE Partial sequence of HIV-1 strain ESS gp41 immunodominant region.

XX KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; ESS.

XX OS Human immunodeficiency virus 1.

XX PN EPI013766-A2.

XX PD 28-JUN-2000.

XX PF 29-NOV-1999; 99EP-00309491.

XX PR 30-NOV-1998; 98US-0110292P.

XX PR 08-FEB-1999; 99US-0119138P.

XX PR 04-NOV-1999; 99US-00433428.

XX PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.

XX PI De Leys R, Zheng J;

XX DR WPI; 2000-402205/35.

XX PT New antigenic peptides and peptide functional derivatives, useful for  
 XX PT detection of antibodies produced in response to human immunodeficiency  
 XX PT virus group O antibodies.

XX PS Example 1; Fig 1; 52pp; English.

XX CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain ESS. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. ESS is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 33 AA;

Query Match 97.8%; Score 91; DB 3; Length 33;  
 Best Local Similarity 93.8%; Pred. No. 1.7e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
 |||||  
 Db 14 LLNSWGCKGRIVCYTS 29

RESULT 9  
 AAW07343  
 ID AAW07343 standard; peptide; 40 AA.

XX

AC AAW07343;  
 XX  
 DT 16-OCT-2003 (revised)  
 DE 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (BSS).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95PR-00002236.  
 XX  
 PA (INERM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;  
 XX  
 DR WPI; 1996-412779/41.  
 DR N-PSDB; AAT44918.  
 XX  
 PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 33; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MWP5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC CCMC I-1544 (BCF02 (BSS)), 1543 (BCF01 (PAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF02 (BSS) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 40 AA;  
 Query Match 97.8%; Score 91; DB 2; Length 40;  
 Best Local Similarity 93.8%; Pred. No. 2.1e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLNSWGCKGRIVCYTS 16  
 Db 19 LLNSWGCKGRIVCYTS 34  
 RESULT 10  
 AAW80469  
 ID AAW80469 standard; peptide; 32 AA.  
 XX  
 AC AAW80469;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX  
 KW Group O human immune deficiency virus; HIV; detection; infection.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX  
 PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX  
 PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX  
 DR WPI; 1998-583190/49.  
 XX  
 PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX  
 PS Claim 6; Page 44; 55pp; French.  
 XX  
 CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 32 AA;  
 Query Match 96.8%; Score 90; DB 2; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 2.4e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLNSWGCKGRIVCYTS 16  
 Db 11 LLNSWGCKGRIVCYTS 26  
 RESULT 11  
 AAW80474  
 ID AAW80474 standard; peptide; 17 AA.  
 XX  
 AC AAW80474;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX  
 KW Group O human immune deficiency virus; HIV; detection; infection.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.

```
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX Claim 6; Page 45; 55pp; French.
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 17 AA;
XX Query Match 95.7%; Score 89; DB 2; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-06;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LNSWGCKGRLCVYTS 16
Db 2 LNSWGCKGRLCVYTS 16
||:|||||
RESULT 12
AAW80467
ID AAW80467 standard; peptide; 16 AA.
XX AC AAW80467;
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
XX Synthetic.
XX Human immunodeficiency virus 1.
XX WO9845323-A1.
XX 15-OCT-1998.
XX 06-APR-1998; 98WO-FR000691.
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX Claim 6; Page 43; 55pp; French.
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 22 AA;
XX Query Match 94.6%; Score 88; DB 2; Length 22;
XX Best Local Similarity 93.8%; Pred. No. 3.3e-06;
XX Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLSWGCKGRLCVYTS 16
Db 1 LLSWGCKGRLCVYTS 16
||:|||||
RESULT 14
AAW80462
ID AAW80462 standard; peptide; 22 AA.
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 16 AA;
XX Query Match 94.6%; Score 88; DB 2; Length 16;
XX Best Local Similarity 93.8%; Pred. No. 2.4e-06;
XX Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLSWGCKGRLCVYTS 16
Db 1 LLSWGCKGRLCVYTS 16
||:|||||
RESULT 13
AAW80461
ID AAW80461 standard; peptide; 22 AA.
XX AC AAW80461;
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
XX Synthetic.
XX Human immunodeficiency virus 1.
XX WO9845323-A1.
XX 15-OCT-1998.
XX 06-APR-1998; 98WO-FR000691.
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX Claim 6; Page 42; 55pp; French.
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 22 AA;
XX Query Match 94.6%; Score 88; DB 2; Length 22;
XX Best Local Similarity 93.8%; Pred. No. 3.3e-06;
XX Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLSWGCKGRLCVYTS 16
Db 1 LLSWGCKGRLCVYTS 16
||:|||||
RESULT 14
AAW80462
ID AAW80462 standard; peptide; 22 AA.
```

```

XX AAW80462;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOPI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 42; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 22 AA;

Query Match 93.5%; Score 87; DB 2; Length 22;
Best Local Similarity 93.8%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||:|||||
Db 1 LLQSWGCKGRLVCYTS 16
   |||:|||||

Search completed: May 7, 2004, 17:42:44
Job time : 30.4538 secs

XX PF 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOPI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 42; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 22 AA;

Query Match 93.5%; Score 87; DB 2; Length 22;
Best Local Similarity 93.8%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||:|||||
Db 1 LLQSWGCKGRLVCYTS 16
   |||:|||||

Search completed: May 7, 2004, 17:42:44
Job time : 30.4538 secs

XX AAW80463;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOPI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 42; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 22 AA;

Query Match 94.6%; Score 88; DB 2; Length 22;
Best Local Similarity 93.8%; Pred. No. 3.3e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||:|||||
Db 1 LLSSWGCKGRLVCYTS 16
   |||:|||||

RESULT 15
AAW80463
ID AAW80463 standard; peptide; 22 AA.
XX
XX AC AAW80463;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.

```

THIS PAGE BLANK (USP10)

US-09-433-428D-15

Query Match 93.5%; Score 87; DB 3; Length 33;  
Best Local Similarity 93.8%; Pred. No. 7.5e-07;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
   | | | | | | | | | | | | | | | |  
Db 14 LLLNLWGCKGRLVCYTS 29

Search completed: May 7, 2004, 17:53:26  
Job time : 8.80672 secs





;; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
;; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESS: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/894,699  
;; FILING DATE: 01-DEC-1997

;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR96/00294  
;; FILING DATE: 26-FEB-1996

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614

;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220

;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide  
;; US-08-894-699-36

Query Match 97.8%; Score 91; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 2.2e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRIVCYTS 16  
Db 19 LLNSWGCKGRIVCYTS 34

RESULT 9  
US-09-444-410-36  
; Sequence 36, Application US/09444410  
; Patent No. 6270975

;; GENERAL INFORMATION:  
;; APPLICANT: SIMON, FRANCOIS  
;; APPLICANT: SARAGOSTI, SENTOB  
;; APPLICANT: LOUSSET-AJAKA, IBITISSAM  
;; APPLICANT: LY, THOAI-DUONG  
;; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
;; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
;; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESS: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA

;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/444,410  
;; FILING DATE:

;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/894,699  
;; FILING DATE:

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614

;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220

;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide  
;; US-09-444-410-36

Query Match 97.8%; Score 91; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 2.2e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRIVCYTS 16  
Db 19 LLNSWGCKGRIVCYTS 34

RESULT 10

US-09-462-917A-137  
; Sequence 137, Application US/09462917A  
; Patent No. 6511801

;; GENERAL INFORMATION:  
;; APPLICANT: Delaporte, Eric  
;; APPLICANT: Peeters, Martine  
;; APPLICANT: Saman, Eric  
;; APPLICANT: Vanden Haesevelde, Marlen  
;; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
;; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
;; CURRENT APPLICATION NUMBER: US/09/462,917A  
;; CURRENT FILING DATE: 2000-04-03  
;; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
;; PRIOR FILING DATE: 1998-07-20  
;; NUMBER OF SEQ ID NOS: 152  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 137  
;; LENGTH: 23  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(23)

US-09-462-917A-137

Query Match 93.5%; Score 87; DB 4; Length 23;  
Best Local Similarity 93.8%; Pred. No. 5.2e-07;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRIVCYTS 16  
Db 19 LLNSWGCKGRIVCYTS 34

STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 100.0%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16  
Db 19 LLNSWGCKGRILVCYTS 34

RESULT 6  
US-09-444-410-68  
Sequence 68, Application US/09444410  
Patent No. 6270975  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 100.0%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16  
Db 19 LLNSWGCKGRILVCYTS 34

RESULT 7  
US-09-433-428D-30  
Sequence 30, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433,428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-30

Query Match 97.8%; Score 91; DB 3; Length 33;  
Best Local Similarity 93.8%; Pred. No. 1.8e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16  
Db 14 LLNSWGCKGRILVCYTS 29

RESULT 8  
US-08-894-699-36  
Sequence 36, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE

Qy 1 LLNSWGCKGRLVCYTS 16  
| | | | | | | | | | | | | | | | | |  
Db 14 LLNSWGCKGRLVCYTS 29

## RESULT 3

US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESS: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-39

Query Match 100.0%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
| | | | | | | | | | | | | | | | | |  
Db 19 LLNSWGCKGRLVCYTS 34

## RESULT 4

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-68

Query Match 100.0%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
| | | | | | | | | | | | | | | | | |  
Db 19 LLNSWGCKGRLVCYTS 34

## RESULT 5

US-08-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 8.80672 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-10  
Perfect score: 93  
Sequence: 1 LLNSWGCKGRLVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/protdata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/protdata/2/iaa/5A\_COMB.pep:\*  
3: /cgn2\_6/protdata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/protdata/2/iaa/6A\_COMB.pep:\*  
5: /cgn2\_6/protdata/2/iaa/6A\_COMB.pep:\*  
6: /cgn2\_6/protdata/2/iaa/6A\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	33	3	US-09-433-428D-6
2	93	100.0	33	3	US-09-433-428D-25
3	93	100.0	40	3	US-08-894-699-39
4	93	100.0	40	3	US-08-894-699-68
5	93	100.0	40	3	US-09-444-410-39
6	93	100.0	40	3	US-09-444-410-68
7	91	97.8	40	3	US-09-433-428D-30
8	91	97.8	40	3	US-08-894-699-36
9	91	97.8	40	3	US-09-444-410-36
10	87	93.5	23	4	US-09-462-917A-137
11	87	93.5	24	4	US-09-462-917A-92
12	87	93.5	24	4	US-09-462-917A-138
13	87	93.5	33	3	US-09-433-428D-4
14	87	93.5	33	3	US-09-433-428D-7
15	87	93.5	33	3	US-09-433-428D-15
16	87	93.5	41	3	US-08-894-699-69
17	87	93.5	41	3	US-09-444-410-69
18	87	93.5	42	3	US-08-894-699-66
19	87	93.5	42	3	US-09-444-410-66
20	87	93.5	116	4	US-09-462-917A-20
21	87	93.5	117	4	US-09-462-917A-6
22	87	93.5	715	4	US-09-462-917A-134
23	86	92.5	23	3	US-09-433-428D-59
24	86	92.5	23	3	US-09-433-428D-60
25	86	92.5	23	3	US-09-433-428D-61
26	86	92.5	23	4	US-09-462-917A-95
27	86	92.5	23	4	US-09-462-917A-99

28	86	92.5	28	3	US-09-433-428D-62	Sequence 62, Appl
29	86	92.5	30	3	US-09-433-428D-63	Sequence 63, Appl
30	86	92.5	30	3	US-09-433-428D-69	Sequence 69, Appl
31	86	92.5	33	3	US-09-433-428D-5	Sequence 5, Appl
32	86	92.5	33	3	US-09-433-428D-8	Sequence 8, Appl
33	86	92.5	33	3	US-09-433-428D-14	Sequence 14, Appl
34	86	92.5	33	3	US-09-433-428D-16	Sequence 16, Appl
35	86	92.5	33	3	US-09-433-428D-29	Sequence 29, Appl
36	86	92.5	35	3	US-09-433-428D-64	Sequence 64, Appl
37	86	92.5	41	3	US-08-894-699-67	Sequence 67, Appl
38	86	92.5	41	3	US-09-444-410-67	Sequence 67, Appl
39	86	92.5	113	4	US-09-462-917A-28	Sequence 28, Appl
40	86	92.5	113	4	US-09-462-917A-40	Sequence 40, Appl
41	86	92.5	115	4	US-09-462-917A-24	Sequence 24, Appl
42	86	92.5	149	3	US-09-433-428D-67	Sequence 67, Appl
43	86	92.5	215	2	US-08-912-129A-58	Sequence 58, Appl
44	86	92.5	220	3	US-09-433-428D-66	Sequence 66, Appl
45	86	92.5	245	2	US-08-912-129A-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1  
US-09-433-428D-6  
; Sequence 6, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-6

Query Match 100.0%; Score 93; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 8.9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
Db 14 LLNSWGCKGRLVCYTS 29

RESULT 2  
US-09-433-428D-25  
; Sequence 25, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-25

Query Match 100.0%; Score 93; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 8.9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 7, 2004, 17:49:21 ; Search time 22.3866 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362a-10

Perfect score: 93  
Sequence: 1 LINSWCKKRLVCTYS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	93.5	23	14	US-10-320-786-137
2	87	93.5	24	14	US-10-320-786-137
3	87	93.5	24	14	US-10-320-786-138
4	87	93.5	116	14	US-10-320-786-20
5	87	93.5	117	14	US-10-320-786-6
6	87	93.5	715	14	US-10-320-786-134
7	86	92.5	23	14	US-10-320-786-95
8	86	92.5	23	14	US-10-320-786-95
9	86	92.5	113	14	US-10-320-786-28
10	86	92.5	113	14	US-10-320-786-40
11	86	92.5	115	14	US-10-320-786-24
12	86	92.5	215	8	US-08-911-824-58
13	86	92.5	245	8	US-08-911-824-48
14	86	92.5	281	8	US-08-911-824-120
15	86	92.5	373	8	US-08-911-824-52

16	86	92.5	460	8	US-08-911-824-60	Sequence 60, Appl
17	86	92.5	488	8	US-08-911-824-95	Sequence 50, Appl
18	86	92.5	490	8	US-08-911-824-95	Sequence 95, Appl
19	86	92.5	526	8	US-08-911-824-97	Sequence 97, Appl
20	86	92.5	618	8	US-08-911-824-54	Sequence 54, Appl
21	86	92.5	706	8	US-08-911-824-93	Sequence 93, Appl
22	86	92.5	736	8	US-08-911-824-91	Sequence 91, Appl
23	86	92.5	873	8	US-08-911-824-61	Sequence 61, Appl
24	85	91.4	23	14	US-10-320-786-91	Sequence 91, Appl
25	85	91.4	23	14	US-10-320-786-100	Sequence 100, Appl
26	85	91.4	23	14	US-10-320-786-101	Sequence 101, Appl
27	85	91.4	110	14	US-10-320-786-14	Sequence 14, Appl
28	85	91.4	110	14	US-10-320-786-16	Sequence 16, Appl
29	85	91.4	113	14	US-10-320-786-2	Sequence 2, Appl
30	85	91.4	113	14	US-10-320-786-10	Sequence 10, Appl
31	85	91.4	113	14	US-10-320-786-12	Sequence 12, Appl
32	85	91.4	113	14	US-10-320-786-18	Sequence 18, Appl
33	85	90.3	22	14	US-10-059-271-6	Sequence 6, Appl
34	84	90.3	254	14	US-10-059-271-82	Sequence 82, Appl
35	84	90.3	256	14	US-10-059-271-97	Sequence 97, Appl
36	84	90.3	1231	14	US-10-059-271-94	Sequence 94, Appl
37	83	89.2	356	14	US-10-357-400-12	Sequence 12, Appl
38	81	87.1	22	14	US-10-059-271-7	Sequence 7, Appl
39	81	87.1	23	14	US-10-320-786-97	Sequence 97, Appl
40	81	87.1	113	14	US-10-320-786-26	Sequence 26, Appl
41	81	87.1	116	14	US-10-320-786-22	Sequence 22, Appl
42	80	86.0	35	14	US-10-026-741-101	Sequence 101, Appl
43	80	86.0	37	14	US-10-026-741-7	Sequence 7, Appl
44	80	86.0	37	14	US-10-026-741-90	Sequence 90, Appl
45	80	86.0	113	14	US-10-320-786-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-10-320-786-137  
Sequence 137, Application US/10320786  
Publication No. US20030180759A1  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014-1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: US 09/462,917  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 137  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(23)  
OTHER INFORMATION:  
US-10-320-786-137

Query Match 93.5%; Score 87; DB 14; Length 23;  
Best Local Similarity 93.8%; Pred. No. 6; 7e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LINSWCKKRLVCTYS 16  
DB 4 LINSWCKKRLVCTYS 19.

```

RESULT 2
US-10-320-786-92
; Sequence 92, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
; OTHER INFORMATION:
US-10-320-786-92

```

Query Match	93.5%	Score 87	DB 14	Length 24
Best Local Similarity	93.8%	Pred. No. 7e-06		
Matches	15	Conservative	0	Mismatches 1
				Indels 0
				Gaps 0
QY	1	LINSWGCKGRLLVCYTS	16	
Db	4	LINSWGCKGRLLVCYTS	19	

```

RESULT 3
US-10-320-786-138
; Sequence 138, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
; OTHER INFORMATION:
; US-10-320-786-138

```

Query March	93.5%	Score 87	DB 14	Length 24
Best Local Similarity	93.8%	Pred. No. 7e-06		
Matches 15, Conservative	0	Mismatches 1	Indels 0	Gaps 0
QY	1	LNMGCKGRVLYCTS	16	

```

DB              4  L1N1WCKGR1VVCYTS  19

RESULT 4
US-10-320-786-20
; Sequence 20, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS-014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320.786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 116
; TYPE: prt
; ORGANISM: Human
US-10-320-786-20

```

Query Match	93.5%	Score 87;	DB 14;	Length 116;
Best Local Similarity	93.8%	Pred. No. 2.9e-05;		
Matches	15;	Conservative	0;	Mismatches 1;
			Indels	0;
			Gaps	0;
QY	1	LNLSMGCKGRIVCYTS	16	
Db	39	LNLSMGCKGRIVCYTS	54	

```

? RESULT 5
? US-10-320-786-6
? Sequence 6, Application US/10320786
? Publication No. US20030180759A1
? GENERAL INFORMATION:
? APPLICANT: Delaporte, Eric
? APPLICANT: Peeters, Martine
? APPLICANT: Sman, Eric
? APPLICANT: Vanden Haesevelde, Marleen
? TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
? FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
? CURRENT FILING DATE: 2002-12-16
? PRIOR APPLICATION NUMBER: US 09/462,917
? PRIOR FILING DATE: 2000-04-03
? PRIOR APPLICATION NUMBER: PCT/EP98/04522
? PRIOR FILING DATE: 1998-07-20
? NUMBER OF SEQ ID NOS: 152
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 6
? LENGTH: 117
? TYPE: PRT
? ORGANISM: Human
? US-10-320-786-6

```

Query Match	93.5%	Score 87	DB 14	Length 117
Best Local Similarity	93.8%	Pred. No. 2.9e-05		
Matches 15	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	LLNSGCKGRIVCYTS	16	
Db	39	LLNLIMGCKGRIVCYTS	54	

RESULT 6  
US-10-320-786-134

```
; Sequence 134, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-134

Query Match          93.5%; Score 87; DB 14; Length 715;
Best Local Similarity 93.8%; Pred. No. 0.00015;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 LINSWGCKGRLVCYTS 16
      ||| ||||| ||||| |||||
DB      628 LNLWGCKGRLVCYTS 643

RESULT 7
US-10-320-786-95
; Sequence 95, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-95

Query Match          92.5%; Score 86; DB 14; Length 23;
Best Local Similarity 87.5%; Pred. No. 9.4e-06;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 LINSWGCKGRLVCYTS 16
      ||| ||||| ||||| |||||
DB      4 LNLWGCKGRLVCYTS 19

RESULT 8
US-10-320-786-99
; Sequence 99, Application US/10320786
```

```
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-99

Query Match          92.5%; Score 86; DB 14; Length 23;
Best Local Similarity 87.5%; Pred. No. 9.4e-06;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 LINSWGCKGRLVCYTS 16
      ||| ||||| ||||| |||||
DB      4 LNLWGCKGRLVCYTS 19

RESULT 9
US-10-320-786-28
; Sequence 28, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-28

Query Match          92.5%; Score 86; DB 14; Length 113;
Best Local Similarity 87.5%; Pred. No. 3.9e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 LINSWGCKGRLVCYTS 16
      ||| ||||| ||||| |||||
DB      39 LNLWGCKGRLVCYTS 54

RESULT 10
US-10-320-786-40
; Sequence 40, Application US/10320786
; Publication No. US20030180759A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Misc feature
; LOCATION: (1)-(113)
; OTHER INFORMATION: Xaa = unknown
US-10-320-786-40
```

```
Query Match          92.5%; Score 86; DB 14; Length 113;
Best Local Similarity 87.5%; Pred. No. 3.9e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LINSWGCKGRIVCYTS 16
      ||| |||||:||||
Db      39 LNLWGCKGRILCYTS 54
```

```
RESULT 11
US-10-320-786-24
; Sequence 24, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-24
```

```
Query Match          92.5%; Score 86; DB 14; Length 115;
Best Local Similarity 87.5%; Pred. No. 4e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LINSWGCKGRIVCYTS 16
      ||| |||||:||||
Db      39 LNLWGCKGRILCYTS 54
```

```
RESULT 12
US-08-911-824-58
; Sequence 58, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL
US-08-911-824-58
```

```
Query Match          92.5%; Score 86; DB 8; Length 215;
Best Local Similarity 87.5%; Pred. No. 7e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LINSWGCKGRIVCYTS 16
      ||| |||||:||||
Db      127 LNLWGCKGRILCYTS 142
```

```
RESULT 13
US-08-911-824-48
; Sequence 48, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL
US-08-911-824-48
```

```
Query Match          92.5%; Score 86; DB 8; Length 245;
Best Local Similarity 87.5%; Pred. No. 7.8e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LINSWGCKGRIVCYTS 16
      ||| |||||:||||
Db      127 LNLWGCKGRILCYTS 142
```

```
RESULT 14
US-08-911-824-120
; Sequence 120, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
```



Query Match	92.5%	Score 86	DB 8	Length 281
Best Local Similarity	87.5%	Pred. No. 8.9e-05		
Matches 14	Conservative	1	Mismatches 1	Indels 0
			Gaps	0

RESULT 15  
US-08-911-824-52  
; Sequence 52, Application US/08911824

Query Match	92.5%	Score 86	DB 8	Length 373
Best Local Similarity	87.5%	Pred. No. 0.00011		
Matches 14; Conservative	1	Mismatches 1	Indels 0	Gaps 0

Search completed: May 7, 2004, 18:29:21  
Job time : 22.3866 secs

THIS PAGE BLANK (CSPTC)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 6.72269 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-10

Perfect score: 93

Sequence: 1 LLNSWGCKGRLVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	86.0	877	2 A49197	envelope protein p
2	79	84.9	863	2 A53034	gag polyprotein -
3	75	80.6	104	2 S2930	GP41 ENV protein -
4	64	68.8	855	2 A45713	Env transmembrane
5	63	67.7	357	2 S21990	envelope protein g
6	63	67.7	358	2 S22002	envelope protein g
7	63	67.7	854	1 VCLJST	env polyprotein pr
8	63	67.7	859	1 VCLJST	env polyprotein pr
9	62	66.7	358	2 S22000	envelope protein g
10	62	66.7	358	2 S70417	envelope protein g
11	60	64.5	151	2 S30458	env protein - huma
12	60	64.5	151	2 S30459	env protein - huma
13	60	64.5	151	2 S30458	env protein - huma
14	60	64.5	151	2 S30453	env protein - huma
15	60	64.5	151	2 S30452	env protein - huma
16	60	64.5	151	2 S30450	env protein - huma
17	60	64.5	151	2 S30451	env protein - huma
18	60	64.5	151	2 S30457	env protein - huma
19	60	64.5	151	2 S30456	env protein - huma
20	60	64.5	151	2 S30455	env protein - huma
21	60	64.5	151	2 S30454	env protein - huma
22	60	64.5	357	2 S22006	envelope protein g
23	60	64.5	357	2 S21994	envelope protein g
24	60	64.5	357	2 S22004	envelope protein g
25	60	64.5	357	2 S21996	envelope protein g
26	60	64.5	357	2 S21992	envelope protein g
27	60	64.5	358	2 S21998	envelope protein g
28	60	64.5	366	2 B41565	env polyprotein -
29	60	64.5	443	2 C41621	env polyprotein P

30	60	64.5	445	2 A41621	env polyprotein M
31	60	64.5	454	2 B41621	env polyprotein D
32	60	64.5	712	1 VCLJSA	env polyprotein pr
33	60	64.5	843	1 H44001	env polyprotein pr
34	60	64.5	847	2 T09448	envelope glycoprot
35	60	64.5	847	2 S13289	env protein - huma
36	60	64.5	851	2 S12159	env protein - huma
37	60	64.5	852	1 VCLJBR	env polyprotein -
38	60	64.5	852	1 VCLJGG	env polyprotein pr
39	60	64.5	852	2 T12016	envelope glycoprot
40	60	64.5	853	2 S54384	envelope polyproce
41	60	64.5	854	2 S13288	env protein huma
42	60	64.5	855	1 VCLJAZ	env polyprotein pr
43	60	64.5	855	1 VCLJZR	env polyprotein pr
44	60	64.5	856	1 VCLJH3	env polyprotein pr
45	60	64.5	856	1 VCLJVL	env polyprotein pr

ALIGNMENTS

RESULT 1

S49197

envelope protein precursor - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 24-Jul-1997 #text\_change 26-Aug-1999

C:Accession: S49197

R:Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy,

submitted to the EMBL Data Library, July 1994

A:Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: defi

A:Reference number: S49197

A:Accession: S49197

A:Molecule type: DNA

A:Residues: 1-877 <CHA>

A:Cross-references: EMBL:X80020; NID:G510516; PIDN:CAA56323.1; PID:G510517

A:Experimental source: isolate VAU

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-535/Product: coat protein gp120 #status predicted <CPI>

F:536-877/Product: coat protein gp41 #status predicted <CP2>

F:698-716/Domain: transmembrane #status predicted <TMN>

F:59,88,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404,

Query Match 86.0%; Score 80; DB 2; Length 877;

Best Local Similarity 81.2%; Pred. No. 0.00021;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16

Db 607 LLNSWGCKNRLICYTS 622

RESULT 2

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborcht, B.; van der Groen,

J. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African h

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-863 <VAN>

A:Cross-references: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 84.9%; Score 79; DB 2; Length 863;

Best Local Similarity 81.2%; Pred. No. 0.00029;

```
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTS 16
   ||: ||||: |||||
Db 594 LLSWGCKGRLVCYTS 609

RESULT 3
S52930
Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 06-Jun-1995 #sequence_revision 21-Jul-1995 #text_change 26-Aug-1999
C:Accession: S52930
R:Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chanaret, S.; Tabary, T.; Montagnier, L.; de
submitted to the EMBL Data Library, January 1995
A:Description: A novel HIV1-O strain illustrates the diversity of the O group.
A:Reference number: S52929
A:Accession: S52930
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-104 <COH>
A:Cross-references: EMBL:X84328; NID:g695526; PIDN:CAA59066.1; PID:g695527
C:Superfamily: type E retrovirus env polyprotein

Query Match 80.6%; Score 75; DB 2; Length 104;
Best Local Similarity 68.8%; Pred. No. 0.0002;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTS 16
   ||| ||||: ||||
Db 45 LLSWGCKGRLVCYTS 60

RESULT 4
A45713
Env transmembrane protein gp43 - human immunodeficiency virus type 2
C:Species: human immunodeficiency virus type 2, HIV-2
C>Date: 04-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C:Accession: A45713
R:Barnett, S.W.; Quiroga, M.; Werner, A.; Dina, D.; Levy, J.A.
J. Virol. 67, 1006-1014, 1993
A:Title: Distinguishing features of an infectious molecular clone of the highly divergent
A:Reference number: A45713; MUID:93124535; PMID:8419635
A:Accession: A45713
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-855 <BAR>
A:Experimental source: UC1
A:Note: sequence extracted from NCBI backbone (NCBIP:122362)
C:Superfamily: type E retrovirus env polyprotein

Query Match 68.8%; Score 64; DB 2; Length 855;
Best Local Similarity 68.8%; Pred. No. 0.055;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTS 16
   ||||| ||||| ||||:
Db 594 LLSWGCAFRQVCHTT 609

RESULT 5
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 20
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21990; S70423
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Molecule type: DNA
```

```
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70423
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332, 'X', 334-357 <STE2>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C:Superfamily: type E retrovirus env polyprotein

Query Match 67.7%; Score 63; DB 2; Length 357;
Best Local Similarity 62.5%; Pred. No. 0.038;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTS 16
   || ||||| ||||: ||
Db 93 LLSWGCKGRLVCYTS 108

RESULT 6
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 3L
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <STE1>
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, 'X', 335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:g60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 67.7%; Score 63; DB 2; Length 358;
Best Local Similarity 62.5%; Pred. No. 0.038;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTS 16
   || ||||| ||||: ||
Db 94 LLSWGCKGRLVCYTS 109

RESULT 7
VCLJJSI
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: simian immunodeficiency virus SIVcpz
A:Note: host Pan troglodytes (chimpanzee)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09990
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09990
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```

A:Residues: 1-854 <HUE>  
A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CRA36407.1; PID:g58874  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-500/Product: coat protein gp120 #status predicted <CPI>  
F:501-854/Product: coat protein gp41 #status predicted <CP2>  
F:501-517/Domain: transmembrane #status predicted <TM1>  
F:675-693/Domain: transmembrane #status predicted <TM2>  
F:805-821/Domain: transmembrane #status predicted <TM3>  
F:134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,454

Query Match 67.7%; Score 63; DB 1; Length 854;  
Best Local Similarity 56.2%; Pred. No. 0.078;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16  
: ||||| : ||||| :  
Db 582 ILGLMGCSGKAVCYTT 597

RESULT 8  
VCLJST  
env polyprotein precursor - human immunodeficiency virus type 2 (isolate ST)  
N:Alternate names: coat polyprotein  
N:Contains: surface glycoprotein gp120; transmembrane glycoprotein gp41  
C:Species: human immunodeficiency virus type 2, HIV-2  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Sep-1994  
C:Accession: H33943  
R:Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;  
J. Virol. 64, 890-901, 1990  
A:Title: Molecular characterization of an attenuated human immunodeficiency virus type 2  
A:Reference number: A33943; MUID:90112662; PMID:2296086  
A:Accession: H33943  
A:Molecule type: genomic RNA  
A:Residues: 1-859 <KUN>  
A:Cross-references: EMBL:M86924  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-859/Product: env polyprotein #status predicted <ENV>  
F:20-501/Product: surface glycoprotein gp120 #status predicted <SGG>  
F:506-859/Product: transmembrane glycoprotein gp41 #status predicted <TGG>  
F:507-523/Region: hydrophobic  
F:675-694/Domain: transmembrane #status predicted <TMN>  
F:36,69,78,113,119,131,137,145,160,173,186,200,232,235,242,266,272,283,294,304,359,392,4

Query Match 67.7%; Score 63; DB 1; Length 859;  
Best Local Similarity 66.7%; Pred. No. 0.078;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTS 16  
: ||||| : ||||| :  
Db 586 LNSWGCSARQVCYTT 600

RESULT 9  
S22000  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C:Accession: S22000  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined  
A:Reference number: S21990  
A:Accession: S22000  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-358 <STE>  
A:Cross-references: EMBL:X61351  
C:Superfamily: type E retrovirus env polyprotein

Query Match 66.7%; Score 62; DB 2; Length 358;  
Best Local Similarity 66.7%; Pred. No. 0.054;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYT 15  
: ||||| : ||||| :  
Db 94 LLGIWGCGRLLCTT 108

RESULT 10  
S70417  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 3B) (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: patient 3B  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 26-Aug-1999  
C:Accession: S70417  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70417  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STE>  
A:Cross-references: EMBL:X61351; NID:g60184; PIDN:CRA43614.1; PID:g60185  
C:Superfamily: type E retrovirus env polyprotein

Query Match 66.7%; Score 62; DB 2; Length 358;  
Best Local Similarity 66.7%; Pred. No. 0.054;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYT 15  
: ||||| : ||||| :  
Db 94 LLGIWGCGRLLCTT 108

RESULT 11  
S30458  
env protein - human immunodeficiency virus type 2 (fragment)  
C:Species: human immunodeficiency virus type 2, HIV-2  
C:Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 23-Mar-2001  
C:Accession: S30458; S30477  
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
Nature 358, 495-499, 1992  
A:Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in west Africa.  
A:Reference number: S30448; MUID:92350299; PMID:1641038  
A:Accession: S30458  
A:Molecule type: nucleic acid  
A:Residues: 1-151 <GAO>  
A:Cross-references: EMBL:M87142  
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
submitted to the EMBL Data Library, December 1992  
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
A:Reference number: S30460  
A:Accession: S30458  
A:Molecule type: nucleic acid  
A:Residues: 1-151 <GAO>  
A:Cross-references: EMBL:M87142  
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
submitted to the EMBL Data Library, December 1992  
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
A:Reference number: S30460  
A:Accession: S30477  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-151 <GA2>  
A:Cross-references: EMBL:M87141  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein

Query Match 64.5%; Score 60; DB 2; Length 151;  
Best Local Similarity 66.7%; Pred. No. 0.053;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2  LNSWGCKGRLVCYTS 16
      |||||  |||:|
Db      37  LNSWGCAFRQVCHTT 51

RESULT 12
S30459
env protein - human immunodeficiency virus type 2 (fragment)
C:Species: human immunodeficiency virus type 2, HIV-2
C>Date: 02-Dec-1993 #sequence_revision 30-Jan-1998 #text_change 17-Mar-1999
C:Accession: S30459
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
Nature 358, 495-499, 1992
A:Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.
A:Reference number: S30448; MUID:92350299; PMID:1641038
A:Accession: S30459
A>Status: translation not shown
A:Molecule type: nucleic acid
A:Residues: 1-151 <GAO>
A:Cross-references: EMBL:M87143
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein

Query Match      64.5%; Score 60; DB 2; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.053;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2  LNSWGCKGRLVCYTS 16
      |||||  |||:|
Db      37  LNSWGCAFRQVCHTT 51

RESULT 13
S30448
env protein - human immunodeficiency virus type 2 (fragment)
C:Species: human immunodeficiency virus type 2, HIV-2
C>Date: 02-Dec-1993 #sequence_revision 30-Jan-1998 #text_change 17-Mar-1999
C:Accession: S30448; S30449; S30480; S30481
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
Nature 358, 495-499, 1992
A:Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.
A:Reference number: S30448; MUID:92350299; PMID:1641038
A:Accession: S30448
A:Molecule type: nucleic acid
A:Residues: 1-151 <GAO>
A:Cross-references: EMBL:M87069
A:Experimental source: FOENVA13
A:Accession: S30449
A>Status: preliminary; translation not shown
A:Molecule type: nucleic acid
A:Residues: 1-151 <GAW>
A:Cross-references: EMBL:M87071
A:Experimental source: FOENVA3
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A:Reference number: S30460
A:Accession: S30480
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-151 <GA2>
A:Cross-references: EMBL:M87085
A:Accession: S30481
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-151 <GA3>
A:Cross-references: EMBL:M87076
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein
```

```
Query Match      64.5%; Score 60; DB 2; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.053;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 LNSWGCKGRLVCYTS 16

||||| |||:|

Db 37 LNSWGCAFRQVCHTT 51

#### RESULT 14

S30453

env protein - human immunodeficiency virus type 2 (fragment)

C:Species: human immunodeficiency virus type 2, HIV-2

C>Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999

C:Accession: S30453

R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;

Nature 358, 495-499, 1992

A:Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.

A:Reference number: S30448; MUID:92350299; PMID:1641038

A:Accession: S30453

A>Status: translation not shown

A:Molecule type: nucleic acid

A:Residues: 1-151 <GAO>

A:Cross-references: EMBL:M87089

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein

```
Query Match      64.5%; Score 60; DB 2; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.053;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 LNSWGCKGRLVCYTS 16

||||| |||:|

Db 37 LNSWGCAFRQVCHTT 51

#### RESULT 15

S30452

env protein - human immunodeficiency virus type 2 (fragment)

C:Species: human immunodeficiency virus type 2, HIV-2

C>Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999

C:Accession: S30452

R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;

Nature 358, 495-499, 1992

A:Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.

A:Reference number: S30448; MUID:92350299; PMID:1641038

A:Accession: S30452

A>Status: translation not shown

A:Molecule type: nucleic acid

A:Residues: 1-151 <GAO>

A:Cross-references: EMBL:M87075

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein

```
Query Match      64.5%; Score 60; DB 2; Length 151;
Best Local Similarity 66.7%; Pred. No. 0.053;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 LNSWGCKGRLVCYTS 16

||||| |||:|

Db 37 LNSWGCAFRQVCHTT 51

Search completed: May 7, 2004, 17:51:02

Job time : 6.72269 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 4.03361 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-10

Perfect score: 93

Sequence: 1 LLNSWGCKGRGLVCYTS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	65	69.9	857	1 ENV_HV2KR	Q74126 human immun
2	63	67.7	854	1 ENV_SIVCZ	P17281 chimpanzee
3	63	67.7	856	1 ENV_HV2NZ	P05883 human immun
4	60	64.5	712	1 ENV_HV2S2	P32536 human immun
5	60	64.5	843	1 ENV_HV1Y2	P35961 human immun
6	60	64.5	846	1 ENV_HV2SB	P12449 human immun
7	60	64.5	847	1 ENV_HV1S1	P19550 human immun
8	60	64.5	847	1 ENV_HV1W2	P05880 human immun
9	60	64.5	851	1 ENV_HV1B8	P04582 human immun
10	60	64.5	851	1 ENV_HV2D1	P17755 human immun
11	60	64.5	851	1 ENV_HV1BN	P18040 human immun
12	60	64.5	852	1 ENV_HV1S3	P12488 human immun
13	60	64.5	852	1 ENV_HV1S3	P19549 human immun
14	60	64.5	853	1 ENV_HV1MF	P19551 human immun
15	60	64.5	853	1 ENV_HV1Z2	P12487 human immun
16	60	64.5	855	1 ENV_HV1A2	P03378 human immun
17	60	64.5	855	1 ENV_HV1OY	P20888 human immun
18	60	64.5	855	1 ENV_HV1Z6	P04580 human immun
19	60	64.5	856	1 ENV_HV1B1	P03375 human immun
20	60	64.5	856	1 ENV_HV1H2	P04578 human immun
21	60	64.5	856	1 ENV_HV1W	Q70626 human immun
22	60	64.5	856	1 ENV_HV1MN	P05877 human immun
23	60	64.5	856	1 ENV_HV1PV	P03376 human immun
24	60	64.5	856	1 ENV_HV1SC	P05878 human immun
25	60	64.5	856	1 ENV_HV1W1	P31872 human immun
26	60	64.5	858	1 ENV_HV2R0	P04577 human immun
27	60	64.5	859	1 ENV_HV2CA	P24105 human immun
28	60	64.5	859	1 ENV_HV2D2	P15831 human immun
29	60	64.5	859	1 ENV_HV2ST	P20872 human immun
30	60	64.5	860	1 ENV_HV2BE	P18094 human immun
31	60	64.5	861	1 ENV_HV1BR	P03377 human immun
32	60	64.5	865	1 ENV_HV1RH	P04579 human immun
33	60	64.5	867	1 ENV_HV1J3	P12489 human immun

34	60	64.5	868	1 ENV_HV1C4	P05879 human immun
35	60	64.5	885	1 ENV_SIVS4	P12492 simian immu
36	60	64.5	889	1 ENV_SIVSP	P19503 simian immu
37	58	62.4	848	1 ENV_HV1JR	P20871 human immu
38	58	62.4	854	1 ENV_SIVAI	Q02837 simian immu
39	58	62.4	856	1 ENV_HV1H3	P04624 human immu
40	58	62.4	865	1 ENV_SIVAT	P05886 simian immu
41	57	61.3	380	1 ENV_SIVM2	P08910 simian immu
42	57	61.3	856	1 ENV_HV1ZH	P05881 human immu
43	57	61.3	881	1 ENV_SIVMK	P05884 simian immu
44	57	61.3	882	1 ENV_SIVM1	P05885 simian immu
45	56	60.2	846	1 ENV_HV1ND	P18799 human immu

ALIGNMENTS

RESULT 1	ENV_HV2KR	STANDARD;	PRT;	857 AA.
ID	ENV_HV2KR			
AC	Q74126;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 2 (isolate KR) (HIV-2).			
OC	Viruses; Retroviridae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=73484;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kraus G.K., Talbott R., Leavitt M., Luzznick L., Schmidt A.,			
RA	Badal P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;			
RL	Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; U22047; AAA64582.1; -			
DR	InterPro; IPR000328; Env.GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL.	1	19	POTENTIAL.
FT	CHAIN	20	503	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	857	TRANSMEMBRANE GLYCOPROTEIN.
FT	CARBOHYD	36	36	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	69	69	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	78	78	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	113	113	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	121	121	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	142	142	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	159	159	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	198	198	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	264	264	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	281	281	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 857 AA; 98689 MW; F881C675B5746DF CRC64;

Query Match 69.9%; Score 65; DB 1; Length 857;
Best Local Similarity 73.3%; Pred. No. 0.0064;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LNSWGCKGRLYCYTS 16
||||| |||||
Db 584 LNSWGCAFRQVCYIT 598

RESULT 2
ENV_SIVCZ STANDARD; PRT; 854 AA.
AC P17281,
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90259077; PubMed=2188136;
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";
RL Nature 345:356-359(1990).
CC -!- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR PIR; S09990; VCLJ51.
DR HIV; X52154; ENVSCPZ.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 500 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 501 854 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 501 517 POTENTIAL.
FT TRANSMEM 675 693 POTENTIAL.
FT TRANSMEM 805 821 POTENTIAL.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 67.7%; Score 63; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 0.014;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLYCYTS 16
:| ||| | |||
Db 582 ILGLWGCGRAVCYIT 597

RESULT 3
ENV_HV2NZ STANDARD; PRT; 856 AA.
AC P05883;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate NIH-Z) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88320359; PubMed=3261862;
RA Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starcich B.R.,
RA Hall L., Fagnoli K., Jagodzinski L.L., Guo H.-G., Laure F.,
RA Arya S.K., Josephs S.F., Zagury D., Wong-Straal F., Gallo R.C.;
RT "Genetic variability between isolates of human immunodeficiency virus
(HIV) type 2 is comparable to the variability among HIV type 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03654; AAB00761.1; -.
DR HIV; J03654; ENV52NIH2.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 492 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 493 856 TRANSMEMBRANE GLYCOPROTEIN.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).

```



```

FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 98665 MW; A938B0A7E2B881D6 CRC64;

Query Match 67.7%; Score 63; DB 1; Length 856;
Best Local Similarity 73.3%; Pred. No. 0.014;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTS 16
Db 583 LNSWGCAFRQVCHTS 597
||||| | | | |

RESULT 4
ENV HV2S2
ID ENV HV2S2 STANDARD; PRT; 712 AA.
AC P32536;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate ST/24.1CH2) (HIV-2).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=31681;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260861; PubMed=1583738;
RA Mulligan M.J., Yamshchikov G.V., Ritter G.D. Jr., Gao F., Jin M.J.,
Nail C.D., Spies C.P., Hahn B.H., Compans R.W.;
RA "Cytoplasmic domain truncation enhances fusion activity by the
RT exterior glycoprotein complex of human immunodeficiency virus type 2
in selected cell types.";
RL J. Virol. 66:3971-3975 (1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86924; AAA43938.1; -.
DR PIR; A42535; VCLJ54.
DR InterPro; IPR000328; ENV_GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 19
FT CHAIN 20 712 ENV POLYPEPTIDE.

```

```

FT CHAIN 20 495 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 496 712 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 675 694 POTENTIAL.
FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 605 605 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81723 MW; 4EC7F3C8D3C3489 CRC64;

Query Match 64.5%; Score 60; DB 1; Length 712;
Best Local Similarity 66.7%; Pred. No. 0.035;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTS 16
Db 586 LNSWGCAFRQVCHYT 600
||||| | | | |

RESULT 5
ENV HV1Y2
ID ENV HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
Shaw G.M.;
RA "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600 (1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 ON NCBI\_TaxID=11691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90347835; PubMed=2384920;  
 RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;  
 RT "Viral determinants of human immunodeficiency virus type 1 T-cell or  
 macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";  
 RL J. Virol. 64:4390-4398 (1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M65024; AAA45072.1; -;  
 DR PDB; 1OBE; 15-MAY-97.  
 DR HIV; M38428; ENV\$SF162.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 502  
 FT CHAIN 503 847  
 FT DISULFID 53 73  
 FT DISULFID 118 203  
 FT DISULFID 125 194  
 FT DISULFID 130 155  
 FT DISULFID 216 245  
 FT DISULFID 226 237  
 FT DISULFID 294 328  
 FT DISULFID 374 435  
 FT DISULFID 381 408  
 FT CARBOHYD 87 87  
 FT CARBOHYD 135 135  
 FT CARBOHYD 134 134  
 FT CARBOHYD 186 186  
 FT CARBOHYD 195 195  
 FT CARBOHYD 232 232  
 FT CARBOHYD 239 239  
 FT CARBOHYD 260 260  
 FT CARBOHYD 274 274  
 FT CARBOHYD 293 293  
 FT CARBOHYD 299 299  
 FT CARBOHYD 329 329  
 FT CARBOHYD 336 336  
 FT CARBOHYD 352 352  
 FT CARBOHYD 382 382  
 FT CARBOHYD 388 388  
 FT CARBOHYD 392 392  
 FT CARBOHYD 398 398  
 FT CARBOHYD 401 401  
 FT CARBOHYD 438 438  
 FT CARBOHYD 454 454  
 FT CARBOHYD 602 602  
 FT CARBOHYD 607 607  
 FT CARBOHYD 616 616  
 FT CARBOHYD 628 628  
 FT SEQUENCE 847 AA; 96135 MW; 0A90317DFDFFZAB CRC64;  
 Query Match 64.5%; Score 60; DB 1; Length 847;  
 Best Local Similarity 56.2%; Pred. No. 0.041;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 LINSWGCKGRILVCYTS 16  
 Db 583 LLGIWGCCKGLICITTA 598  
 RESULT 8  
 ENV HV1W2 STANDARD; PRT; 847 AA.  
 AC P05860;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 ON NCBI\_TaxID=11705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86235450; PubMed=3012778;  
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
 RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks W.P.;  
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or  
 at risk for AIDS";  
 RL Science 232:1548-1553 (1986).  
 CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
 BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
 WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M12507; AAB12990.1; -;  
 DR HIV; M12507; ENV\$WMJ2.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 501  
 FT CHAIN 502 847  
 FT DISULFID 53 73  
 FT DISULFID 118 202  
 FT DISULFID 125 193  
 FT DISULFID 130 152  
 FT DISULFID 215 244  
 FT DISULFID 225 236  
 FT DISULFID 293 326  
 FT DISULFID 372 435  
 FT DISULFID 379 408  
 FT CARBOHYD 87 87  
 FT CARBOHYD 134 134  
 FT CARBOHYD 140 140  
 FT CARBOHYD 151 151  
 FT CARBOHYD 155 155  
 FT CARBOHYD 183 183  
 FT CARBOHYD 184 184  
 FT CARBOHYD 194 194  
 FT CARBOHYD 231 231  
 FT CARBOHYD 238 238  
 FT CARBOHYD 259 259  
 FT CARBOHYD 273 273  
 FT CARBOHYD 286 286  
 FT SEQUENCE 501 AA; 96135 MW; 0A90317DFDFFZAB CRC64;  
 Query Match 64.5%; Score 60; DB 1; Length 847;  
 Best Local Similarity 56.2%; Pred. No. 0.041;







FT DISULFID 219 248 BY SIMILARITY. (POTENTIAL).  
 FT DISULFID 229 240 BY SIMILARITY. (POTENTIAL).  
 FT DISULFID 297 331 BY SIMILARITY. (POTENTIAL).  
 FT DISULFID 377 439 BY SIMILARITY. (POTENTIAL).  
 FT DISULFID 384 412 BY SIMILARITY. (POTENTIAL).  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 852 AA; 96663 MW; EE78BF8D23C9910D CRC64;  
 Query Match 64.5%; Score 60; DB 1; Length 852;  
 Best Local Similarity 56.2%; Pred. No. 0.042;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 LLNSWGCKGRVLCYTS 16  
 DB 588 LLGIWCGSGKLCICTTT 603  
 RESULT 14  
 ID ENV\_HV1MF STANDARD; PRT; 853 AA.  
 AC P19551;  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).  
 OC Viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11704;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317877; PubMed=1695254;  
 RA Stevenson M., Haggerty S., Lamonia C., Mann A.M., Meier C., Wasiaak A.;  
 RT "Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytolysis";  
 RT J. Virol. 64:3792-3803 (1990).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M33943; AAA44850.1; -.  
 DR PDB; IAIK; 16-JUN-97.  
 DR HIV; M33943; ENVSMFA.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR00777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 3D-structure. 1 30  
 SIGNAL 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.  
 FT CHAIN 54 74 BY SIMILARITY.  
 FT DISULFID 119 203 BY SIMILARITY.  
 FT DISULFID 126 194 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 216 245 BY SIMILARITY.  
 FT DISULFID 226 237 BY SIMILARITY.  
 FT DISULFID 294 329 BY SIMILARITY.  
 FT DISULFID 376 443 BY SIMILARITY.  
 FT DISULFID 383 416 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 672 672 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 814 814 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 853 AA; 96912 MW; 337F993B6F22ABA CRC64;  
 Query Match 64.5%; Score 60; DB 1; Length 853;  
 Best Local Similarity 56.2%; Pred. No. 0.042;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 LLNSWGCKGRVLCYTS 16  
 DB 590 LLGIWCGSGKLCICTTA 605  
 RESULT 15  
 ID ENV\_HV122 STANDARD; PRT; 853 AA.  
 AC P12487;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].



GN ENV.  
OS Human immunodeficiency virus type 1 (Z2/CDC-234 isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11683;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Theodore T., Buckler-White A.;  
RL Submitted (NOV-1988) to the HIV data bank.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M22639; AAA45370.1; -;  
DR PIR; S54384; S54384.  
DR HIV; M22639; ENV\$Z226.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW Signal.  
KW SIGNAL.  
FT SIGNAL 1 31 BY SIMILARITY.  
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 206 BY SIMILARITY.  
FT DISULFID 125 197 BY SIMILARITY.  
FT DISULFID 130 154 BY SIMILARITY.  
FT DISULFID 219 248 BY SIMILARITY.  
FT DISULFID 229 240 BY SIMILARITY.  
FT DISULFID 297 330 BY SIMILARITY.  
FT DISULFID 376 442 BY SIMILARITY.  
FT DISULFID 383 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 853 AA; 97043 MW; 849B0B8CAFF7008 CRC64;  
Query Match 64.5%; Score 60; DB 1; Length 853;  
Best Local Similarity 56.2%; Pred. No. 0.042;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16  
Db 589 LLGIWGCKGLICTTT 604  
Search completed: May 7, 2004, 17:43:56  
Job time : 4.03361 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 20.3025 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-10

Perfect score: 93  
Sequence: 1 LLNSWGCKGRIVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_xvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	93	100.0	116 15 Q7ZJN9	Q7ZJN9 human immun
2	93	100.0	216 15 Q9IEC5	Q9IEC5 human immun
3	93	100.0	219 15 Q9IEB6	Q9IEB6 human immun
4	93	100.0	890 15 Q8Q7G2	Q8Q7G2 human immun
5	91	97.8	219 15 Q9IEC8	Q9IEC8 human immun
6	90	96.8	130 15 Q9IHU9	Q9IHU9 human immun
7	90	96.8	872 15 Q8Q7H0	Q8Q7H0 human immun
8	90	96.8	882 15 Q8Q7F9	Q8Q7F9 human immun
9	90	96.8	887 15 Q8Q7H6	Q8Q7H6 human immun
10	90	96.8	887 15 Q8Q7G9	Q8Q7G9 human immun
11	89	95.7	135 15 Q9DQL9	Q9DQL9 human immun
12	89	95.7	242 15 Q9IE31	Q9IE31 human immun
13	87	93.5	115 15 Q7ZJN8	Q7ZJN8 human immun
14	87	93.5	116 15 Q40459	Q40459 human immun
15	87	93.5	118 15 Q40451	Q40451 human immun
16	87	93.5	120 15 Q9IHU2	Q9IHU2 human immun

17	87	93.5	131	15	Q9WR05	Q9wr05 human immun
18	87	93.5	131	15	Q9IHU0	Q9ihu0 human immun
19	87	93.5	137	15	Q9IHV5	Q9ihv5 human immun
20	87	93.5	146	15	Q9WRV2	Q9wrv2 human immun
21	87	93.5	213	15	Q9IEC3	Q9iec3 human immun
22	87	93.5	214	15	Q9DIK3	Q9dik3 human immun
23	87	93.5	214	15	Q9IE96	Q9ie96 human immun
24	87	93.5	216	15	Q9IEA5	Q9iea5 human immun
25	87	93.5	342	15	O11942	O11942 human immun
26	87	93.5	532	15	Q9IEB9	Q9ieb9 human immun
27	87	93.5	544	15	Q9IED9	Q9ied9 human immun
28	87	93.5	548	15	Q9IED6	Q9ied6 human immun
29	87	93.5	867	15	Q8Q7G8	Q8q7g8 human immun
30	87	93.5	871	15	O57073	O57073 human immun
31	87	93.5	871	15	O57074	O57074 human immun
32	87	93.5	871	15	Q8Q7I2	Q8q7i2 human immun
33	87	93.5	872	15	Q900Y5	Q900y5 human immun
34	87	93.5	872	15	O57072	O57072 human immun
35	87	93.5	900	15	Q9QNZ8	Q9qnz8 human immun
36	86	92.5	114	15	O40448	O40448 human immun
37	86	92.5	116	15	O11941	O11941 human immun
38	86	92.5	116	15	O40458	O40458 human immun
39	86	92.5	116	15	O40449	O40449 human immun
40	86	92.5	117	15	Q9Q6F2	Q9q6f2 human immun
41	86	92.5	124	15	Q9IHU7	Q9ihu7 human immun
42	86	92.5	132	15	Q9IHV0	Q9ihv0 human immun
43	86	92.5	134	15	Q9IHV4	Q9ihv4 human immun
44	86	92.5	155	15	Q8J3Q4	Q8j3q4 human immun
45	86	92.5	158	15	Q8J3N6	Q8j3n6 human immun

#### ALIGNMENTS

#### RESULT 1

Q7ZJN9 PRELIMINARY; PRT; 116 AA.  
AC Q7ZJN9;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_taxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=08692A;  
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
(Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
integrase, and env gp41.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214120; AA061840.1; -  
DR GO; GO:0019031; C:viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 116 116  
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match 100.0%; Score 93; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRIVCYTS 16  
| | | | | | | | | | | | | | | | | | | | | |  
Db 32 LLNSWGCKGRIVCYTS 47

#### RESULT 2

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5

```
AC Q9IEC5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF07;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236394; CAB96243.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25027 MW; 413A95BBF1B4FC9A CRC64;

Query Match 100.0%; Score 93; DB 15; Length 216;
Best Local Similarity 100.0%; Pred. No. 9e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
| | | | | | | | | | | | | | | |
Db 47 LLNSWGCKGRLVCYTS 62

RESULT 3
Q9IEB6 PRELIMINARY; PRT; 219 AA.
AC Q9IEB6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF14;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236404; CAB96252.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 100.0%; Score 93; DB 15; Length 219;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
| | | | | | | | | | | | | | | |
Db 57 LLNSWGCKGRLVCYTS 72
```

```
RESULT 4
Q8Q7G2 PRELIMINARY; PRT; 890 AA.
AC Q8Q7G2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97US08692A;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Nansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282 (2002).
DR EMBL; AF383259; AAL98881.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 890 AA; 100610 MW; BB816BC5C45EE23 CRC64;

Query Match 100.0%; Score 93; DB 15; Length 890;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRLVCYTS 16
| | | | | | | | | | | | | | | |
Db 620 LLNSWGCKGRLVCYTS 635

RESULT 5
Q9IEC8 PRELIMINARY; PRT; 219 AA.
AC Q9IEC8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF02;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236391; CAB96240.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;
```

```
Query Match          97.8%; Score 91; DB 15; Length 219;
Best Local Similarity 93.8%; Pred. No. 2e-07;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||||:|||||
DB 56 LLNSWGCKGRIVCYTS 71

RESULT 6
Q91HU9          PRELIMINARY; PRT; 130 AA.
AC Q91HU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CW798;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,
RA Phienazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229235; AAF1912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match          96.8%; Score 90; DB 15; Length 130;
Best Local Similarity 93.8%; Pred. No. 1.8e-07;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||||:|||||
DB 40 LLNSWGCKGRIVCYTS 55

RESULT 7
Q8Q7H0          PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CWA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383251; AAL98873.1; -.

Query Match          96.8%; Score 90; DB 15; Length 882;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||||:|||||
DB 613 LLNSWGCKGRIVCYTA 628

RESULT 8
Q8Q7F9          PRELIMINARY; PRT; 882 AA.
AC Q8Q7F9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CWA100;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383262; AAL98884.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 627BCA5D0C09C6D4 CRC64;

Query Match          96.8%; Score 90; DB 15; Length 882;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTS 16
   |||||:|||||
DB 613 LLNSWGCKGRIVCYTA 628

RESULT 9
Q8Q7H6          PRELIMINARY; PRT; 887 AA.
AC Q8Q7H6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
```

```

OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CNAB124;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383245; AAL98867.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 887 AA; 99366 MW; E210F1E3F7B2474D CRC64;

Query Match          96.8%; Score 90; DB 15; Length 887;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16
   |||||:|||||
Db 618 LLNSWGCKGRILVCYTS 633

RESULT 10
QY Q8Q7G9 PRELIMINARY; PRT; 887 AA.
AC Q8Q7G9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
DE Phylogenetic Clusters.";
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CNAB141;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383252; AAL98874.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 887 AA; 99559 MW; 88954F4ED76A314A CRC64;

Query Match          96.8%; Score 90; DB 15; Length 887;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16
   |||||:|||||
Db 618 LLNSWGCKGRILVCYTS 633

RESULT 11
QY Q9DQL9 PRELIMINARY; PRT; 135 AA.
AC Q9DQL9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20584646; PubMed=11153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,
RA Vadillo J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).
DR EMBL; AF255939; AAG36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match          95.7%; Score 89; DB 15; Length 135;
Best Local Similarity 87.5%; Pred. No. 2.7e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRILVCYTS 16
   |||||:|||||
Db 48 LLNSWGCKGRILVCYTS 63

RESULT 12
QY Q9IE31 PRELIMINARY; PRT; 242 AA.
AC Q9IE31;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF112;
RA Roques P., Robertson D., Sousquiere S., Damond F., Mauchere P.,
RA Depienne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243366; CAB96336.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 242
SQ SEQUENCE 242 AA; 27539 MW; DE8A73DF0E8A6FD7 CRC64;

Query Match          95.7%; Score 89; DB 15; Length 242;
Best Local Similarity 87.5%; Pred. No. 4.8e-07;
```

Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRVLCVYTS 16  
|||||:|:|  
Db 64 LLNSWGCKGRVLCVYTS 79

## RESULT 13

ID Q7ZJN8 PRELIMINARY; PRT; 115 AA.  
AC Q7ZJN8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
EN Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3012;  
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
(Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
integrase, and env gp41.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214121; AA061841.1; -  
DR GO; GO:0019031; C:viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 115  
SQ SEQUENCE 115 AA; 13737 MW; 08E0C0D481E737F9 CRC64;

Query Match 93.5%; Score 87; DB 15; Length 115;

Best Local Similarity 93.8%; Pred. No. 5e-07; Length 115;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRVLCVYTS 16  
|||||:|:|  
Db 32 LLNSWGCKGRVLCVYTS 47

## RESULT 14

ID O40459 PRELIMINARY; PRT; 116 AA.  
AC O40459;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
EN Envelope transmembrane glycoprotein (Fragment).  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=group O;  
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;  
RT "Molecular characterization of envelope transmembrane glycoprotein of  
14 new human immunodeficiency virus type 1 group O strains from  
different African countries.";  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y09775; CAA70914.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 13975 MW; 12B3D0D0D2A1AD32 CRC64;

Query Match 93.5%; Score 87; DB 15; Length 116;

Best Local Similarity 93.8%; Pred. No. 5.1e-07;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRVLCVYTS 16  
|||||:|:|  
Db 39 LLNSWGCKGRVLCVYTS 54

## RESULT 15

ID O40451 PRELIMINARY; PRT; 118 AA.  
AC O40451;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
EN Envelope transmembrane glycoprotein (Fragment).  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=group O;  
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;  
RT "Molecular characterization of envelope transmembrane glycoprotein of  
14 new human immunodeficiency virus type 1 group O strains from  
different African countries.";  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y09779; CAA70918.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 118  
SQ SEQUENCE 118 AA; 14094 MW; C4491740918DE154 CRC64;

Query Match 93.5%; Score 87; DB 15; Length 118;

Best Local Similarity 93.8%; Pred. No. 5.2e-07;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLNSWGCKGRVLCVYTS 16  
|||||:|:|  
Db 39 LLNSWGCKGRVLCVYTS 54

Search completed: May 7, 2004, 17:49:13

Job time : 20.3025 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 60.9076 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-11

Perfect score: 176

Sequence: 1 ALSTLQNLQQLNSWGCRGLVCYTVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	100.0	32	2 AAW80469	AAW80469 Peptide d
2	170	96.6	32	2 AAW80470	AAW80470 Peptide d
3	168	95.5	40	2 AAW07346	AAW07346 Partial s
4	165	93.8	32	2 AAW80471	AAW80471 Peptide d
5	161	91.5	200	3 AAY77373	AAW77373 HIV-1 gro
6	161	91.5	215	2 AAY09499	AAW09499 HIV-1 gro
7	161	91.5	215	2 AAY06983	AAW06983 Recombina
8	161	91.5	215	3 AAY77374	AAW77374 HIV-1 gro
9	161	91.5	245	2 AAY09493	AAW09493 HIV-1 gro
10	161	91.5	245	2 AAY06977	AAW06977 Recombina
11	161	91.5	245	3 AAY77369	AAW77369 HIV-1 gro
12	161	91.5	281	2 AAY03507	AAW03507 HIV-1 gro
13	161	91.5	373	2 AAY09495	AAW09495 HIV-1 gro
14	161	91.5	373	2 AAY06979	AAW06979 Recombina
15	161	91.5	460	2 AAY09500	AAW09500 HIV-1 gro
16	161	91.5	460	2 AAY06984	AAW06984 Recombina
17	161	91.5	460	3 AAY77375	AAW77375 HIV-1 gro
18	161	91.5	474	3 AAY77371	AAW77371 HIV-1 gro
19	161	91.5	488	2 AAY09504	AAW09504 HIV-1 gro
20	161	91.5	490	2 AAY09494	AAW09494 HIV-1 gro
21	161	91.5	490	2 AAY06978	AAW06978 Recombina
22	161	91.5	490	3 AAY77370	AAW77370 HIV-1 gro
23	161	91.5	526	2 AAY09505	AAW09505 HIV-1 gro
24	161	91.5	618	2 AAY09496	AAW09496 HIV-1 gro
25	161	91.5	618	2 AAY06980	AAW06980 Recombina

26	161	91.5	618	3	AAW77372	AAW77372 HIV-1 gro
27	161	91.5	706	2	AAW09503	AAW09503 HIV-1 gro
28	161	91.5	736	2	AAW09502	AAW09502 HIV-1 gro
29	161	91.5	873	2	AAW09501	AAW09501 HIV-1 gro
30	161	91.5	873	2	AAW06985	AAW06985 Amino aci
31	161	91.5	873	3	AAW77376	AAW77376 HIV-1 gro
32	158	89.8	33	3	AAW12231	AAW12231 Partial s
33	158	89.8	40	2	AAW07352	AAW07352 Partial s
34	158	89.8	40	2	AAW07343	AAW07343 Partial s
35	158	89.8	116	2	AAW05555	AAW05555 HIV-1 gro
36	158	89.8	356	2	AAW03940	AAW03940 GP 41 ant
37	157	89.2	41	2	AAW07353	AAW07353 Partial s
38	157	89.2	113	2	AAW05546	AAW05546 HIV-1 gro
39	157	89.2	117	2	AAW05548	AAW05548 HIV-1 gro
40	156	88.6	33	3	AAW12236	AAW12236 Partial s
41	156	88.6	33	3	AAW12212	AAW12212 Partial s
42	156	88.6	41	2	AAW07351	AAW07351 Partial s
43	156	88.6	110	2	AAW05552	AAW05552 HIV-1 gro
44	156	88.6	715	2	AAW05625	AAW05625 HIV-1 gro
45	155	88.1	104	2	AAW07245	AAW07245 HIV-1 gro

ALIGNMENTS

RESULT 1  
AAW80469  
ID AAW80469 standard; peptide; 32 AA.

AC AAW80469;  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
PD 15-OCT-1998.  
XX  
PF 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
PI -Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX  
WPI; 1998-583190/49.  
XX  
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
XX  
PS Claim 6; Page 44; 55pp; French.  
XX  
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 32 AA;

Query Match 100.0%; Score 176; DB 2; Length 32;

```
Best Local Similarity 100.0%; Pred. No. 5.8e-16;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 1 ALETLNQQLLSWGCGRGLVCYTSVRWNET 32

RESULT 2
AAW80470
ID AAW80470 standard; peptide; 32 AA.
XX
AC AAW80470;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
XX
PN W09845323-A1.
XX
PD 15-OCT-1998.
XX
PF 06-APR-1998; 98WO-FR000691.
XX
PR 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002121.
XX
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
PR New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
PS Claim 6; Page 44; 55pp; French.
XX
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 32 AA;

Query Match 96.6%; Score 170; DB 2; Length 32;
Best Local Similarity 96.9%; Pred. No. 3.6e-15;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 1 ALETLNQQLLSWGCGRGLVCYTSVRWNET 32

RESULT 3
AAW07346
ID AAW07346 standard; peptide; 40 AA.
XX
AC AAW07346;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 03-JUN-1997 (first entry)
XX
KW Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.

Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).
Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
primer; hybridisation; amplification; PCR; polymerase chain reaction;
immunogen; antibody.
Human immunodeficiency virus 1.
W09627013-A1.
06-SEP-1996.
26-FEB-1996; 96WO-FR000294.
27-FEB-1995; 95FR-00002236.
(INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
(ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;
WPI; 1996-412779/41.
N-PSDB; AAT44922.
New strains of HIV-1 group O, related DNA fragments, peptide(s) and
antibodies - useful for diagnosis, screening and typing, or as
immunogens.
Claim 12; Page 34; 71pp; French.
Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
into 2 major groups based on the nucleotide sequences of the envelop gene
(env): group M containing sub-groups A-G, and group O containing the
strains ANT70 and MPE5180. The invention relates to the discovery of
several new strains of HIV-1 which can be placed in group O, based on the
partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
and AAW07329-64). The novel strains have been deposited as retroviruses
CNCM I-1544 (BCF02 (SSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
(BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein
encoded by the env gene. The nucleic acids can be used to detect gp. O
HIV-1 strains by hybridisation or (as primers) by gene amplification.
also for screening and typing of such strains. Peptides encoded by the
nucleic acids can be used as immunogens to raise Ab for detecting gp. O
HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
2003 to standardise OS field)
Sequence 40 AA;

Query Match 95.5%; Score 168; DB 2; Length 40;
Best Local Similarity 90.6%; Pred. No. 8.3e-15;
Matches 29; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 9 ALETLNQQLLSWGCGRGLVCYTSVRWNET 40

RESULT 4
AAW80471
ID AAW80471 standard; peptide; 32 AA.
XX
AC AAW80471;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
```



XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 XX 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 44; 55pp; French.  
 XX AA80459-74 represent synthetic peptides (either linear or cyclised by  
 XX Cys-Cys disulphide bonds). The peptides represent variable sequences  
 XX connected around short highly conserved sequences present in isolates of  
 XX group O human immune deficiency virus (HIV). The peptides are useful as  
 XX immunological reagents for detecting infection by group O human immune  
 XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 XX (Updated on 27-AUG-2003 to correct OS field.)  
 XX SQ Sequence 32 AA;  
 Query Match 93.8%; Score 165; DB 2; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 1.6e-14;  
 Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRMNET 32  
 |||||:|||||:|||||:|||||:|||||  
 DB 1 ALETLLNQQLLDLWGCGRGLVCYTSVRMNET 32  
 |||||:|||||:|||||:|||||:|||||  
 RESULT 5  
 AAY77373  
 ID AAY77373 standard; protein; 200 AA.  
 XX AAY77373;  
 XX 22-MAY-2000 (first entry)  
 XX HIV-1 group O env gp120/gp41 pGO-5 recombinant protein, SEQ ID NO:56.  
 XX HIV-i group O; env; gp120; gp41; glycoprotein; monoclonal antibody;  
 XX immunoassay; positive control; affinity purification; therapeutic;  
 XX Escherichia coli; antigen; synthetic gene construction; mutagen;  
 XX deletion mutation.  
 XX Human immunodeficiency virus 1; group O isolate HAM112.  
 OS Synthetic.  
 XX WO200004383-A2.  
 XX 27-JAN-2000.  
 XX 09-JUL-1999; 99WO-US015469.  
 XX 14-JUL-1998; 98US-00115171.  
 XX (ABBO ) ABBOTT LAB.  
 XX Scheffel JW, Hackett JR, Tyner JD, Hickman RK;  
 XX

DR WPI; 2000-171290/15.  
 DR N-PSDB; AA290284.  
 XX Novel monoclonal antibodies useful as positive control reagent for  
 PT detecting human immunodeficiency virus infections and diagnosing,  
 PT evaluating or prognosing viral disease.  
 XX Example 3; Fig 9; 148pp; English.  
 XX The invention relates to anti-HIV-1 group O monoclonal antibodies, which  
 CC may be used as positive control reagents in immunoassays to detect and  
 CC differentiate HIV-1 infections. The invention also encompasses a  
 CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,  
 CC which has no more than 15% cross reactivity to a corresponding antigen  
 CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of  
 CC using a monoclonal antibody as a positive control reagent in an  
 CC immunoassay for the detection of anti HIV-1 group O antibodies. The  
 CC monoclonal antibodies are useful as positive control reagents in  
 CC immunoassays capable of detecting anti-HIV-1 group O antibodies. Such  
 CC immunoassays involve coupling a monoclonal antibody with HIV group-1  
 CC antigen and detecting the antigen-antibody complex. The monoclonal  
 CC antibodies of the invention would be used to ensure that the reagents  
 CC provided to detect HIV-1 group O antibody were performing properly. The  
 CC monoclonal antibodies may also can be immobilised on a matrix and used  
 CC for affinity purification of specific HIV-1 group O-derived proteins from  
 CC cell cultures or biological tissues. The monoclonal antibodies can also  
 CC be used for generating chimeric antibodies for therapeutic use. Different  
 CC synthetic, recombinant or purified antibodies which identify different  
 CC epitopes of HIV antigens can be used in combination in assay to diagnose,  
 CC evaluate, or prognosticate HIV disease condition. The monoclonal  
 CC antibodies are also useful for differentiating HIV-1 Group O antigens  
 CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent  
 CC recombinant HIV-1 group O env antigens encoded by the synthetic genes  
 CC AA290280-Z90286. The recombinant HIV-1 env proteins contain various  
 CC deletions relative to the native HAM112 isolate env protein (AAY77376).  
 CC The recombinant HIV-1 group O antigens were purified and used to screen  
 CC hybridoma cultures  
 XX SQ Sequence 200 AA;  
 Query Match 91.5%; Score 161; DB 3; Length 200;  
 Best Local Similarity 84.4%; Pred. No. 3.7e-13;  
 Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRMNET 32  
 |||||:|||||:|||||:|||||:|||||  
 DB 102 ALETLLNQQLLDLWGCGRGLVCYTSVRMNET 133  
 |||||:|||||:|||||:|||||:|||||  
 RESULT 6  
 AAY09499  
 ID AAY09499 standard; protein; 215 AA.  
 XX AAY09499;  
 XX 17-OCT-2003 (revised)  
 XX 15-JUL-1999 (first entry)  
 XX HIV-1 Group O env polypeptide pGO-8PL.  
 XX HIV; human immunodeficiency virus; antigen; detection; antibody;  
 XX differentiation; Group O; env; immunogen; immunoassay.  
 XX Human immunodeficiency virus 1.  
 XX WO9909179-A2.  
 XX 25-FEB-1999.  
 XX 17-AUG-1998; 98WO-US017014.  
 XX 15-AUG-1997; 97US-00911824.  
 XX

```

PA (ABBO ) ABBOTT LAB.
XX Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
XX WPI; 1999-190167/16.
DR N-PSDB; AAX56078.
XX
XX New isolated HIV-1 Group O env polypeptides - used for the detection of
PT anti-HIV antibodies and for the production of antibodies for use in
PT detection, purification and therapy.
XX
XX Claim 17; Fig 5; 138pp; English.
XX
XX The present invention describes (A) an isolated HIV-1 Group O env
CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env
CC polypeptide comprising an immunoreactive portion of a polypeptide as in
CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at
CC least one HIV-1 Group M env polypeptide; (5) an antigen construct
CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an
CC antigen construct comprising a first HIV-2 env polypeptide fused to a
CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
CC host cell transformed by an expression vector as in (8); and (10) an
CC immunoassay kit for the detection of antibodies to HIV-1 comprising an
CC antigen construct as in (3)-(6). The antigen constructs can be used for
CC the detection of anti-HIV-1 antibodies in test samples. They can also be
CC used as immunogens to produce antibodies. The antibodies can be used to
CC purify HIV polypeptides, for therapy and for detection of HIV
CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 215 AA;

Query Match 91.5%; Score 161; DB 2; Length 215;
Best Local Similarity 84.4%; Pred. No. 4e-13; Indels 0; Gaps 0;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLQNQLLNLSWGCGRGLVCYTSVRWNET 32
Db 117 ALETLQNQLLNLSWGCGRGLVCYTSVRWNET 148

RESULT 7
AAY06983
ID AAY06983 standard; protein; 215 AA.
XX
XX AAY06983;
XX
XX 06-JUL-1999 (first entry)
XX
XX Recombinant pGO-8PL protein.
XX
XX HIV-1, HIV-2; immobilised capture reagent; capillary action; screening;
XX antibody; assay.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX Key Location/Qualifiers
FH Protein 2..46
FT /note= "gp120 sequence"
FT Peptide 47..245
FT /note= "gp41 sequence"
XX
XX WO9909410-A2.
XX
XX 25-FEB-1999.
XX
XX 07-AUG-1998; 98WO-US016506.
XX

PA (ABBO ) ABBOTT LAB.
XX 15-AUG-1997; 97US-00912129.
XX (ABBO ) ABBOTT LAB.
XX
XX Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
PI Golden AM, Brennan CA, Devare SG;
XX
XX WPI; 1999-190224/16.
DR N-PSDB; AAX37193.
XX
XX New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
PT be used in field assay, requiring no electricity and less specialised
PT equipment.
XX
XX Claim 1; Fig 5; 104pp; English.
XX
XX The invention relates to a rapid assay for simultaneous detection and
CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
CC method comprises (a) contacting the sample with a strip containing at
CC least one immobilised capture reagent per analyte and on which the sample
CC moves from the proximal to the distal end by capillary action, under
CC conditions sufficient to form capture reagent/analyte complexes, and (b)
CC determining the presence of analyte(s) by detecting a visible colour
CC change at the capture reagent site on the strip wherein the capture
CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in
CC AAY06981. The invention is used to screen patients for antibodies to HIV-
CC 1 types O and M, and HIV-2. The invention will be particularly useful in
CC places and situation where equipment and/or electricity is not available.
CC The invention provides a screening method which is faster and requires
CC less equipment than prior art methods. The present sequence represents a
CC amino acid sequence of the recombinant pGO-8PL protein which acts as a
CC capture reagent for HIV-1 group O
XX
XX Sequence 215 AA;

Query Match 91.5%; Score 161; DB 2; Length 215;
Best Local Similarity 84.4%; Pred. No. 4e-13; Indels 0; Gaps 0;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLQNQLLNLSWGCGRGLVCYTSVRWNET 32
Db 117 ALETLQNQLLNLSWGCGRGLVCYTSVRWNET 148

RESULT 8
AAY77374
ID AAY77374 standard; protein; 215 AA.
XX
XX AAY77374;
XX
XX 22-MAY-2000 (first entry)
XX
XX HIV-1 group O env gp120/gp41 pGO-8PL recombinant protein, SEQ ID NO:58.
XX
XX HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;
XX immunoassay; positive control; affinity purification; therapeutic;
XX Escherichia coli; antigen; synthetic gene construction; muten;
XX deletion mutation.
XX
XX Human immunodeficiency virus 1; group O isolate HAM112.
XX Synthetic.
XX
XX WO200004383-A2.
XX
XX 27-JAN-2000.
XX
XX 09-JUL-1999; 99WO-US015469.
XX
XX 14-JUL-1998; 98US-00115171.
XX
XX (ABBO ) ABBOTT LAB.

```

XX Scheffel JW, Hackett JR, Tyner JD, Hickman RK;  
 XX WPI; 2000-171290/15.  
 XX N-PSDB; AAZ90285.  
 XX Novel monoclonal antibodies useful as positive control reagent for  
 PT detecting human immunodeficiency virus infections and diagnosing,  
 PT evaluating or prognosing viral disease.  
 XX Example 3; Fig 5; 148pp; English.  
 XX The invention relates to anti-HIV-1 group O monoclonal antibodies, which  
 CC may be used as positive control reagents in immunoassays to detect and  
 CC differentiate HIV-1 infections. The invention also encompasses a  
 CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,  
 CC which has no more than 15% cross reactivity to a corresponding antigen  
 CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of  
 CC using a monoclonal antibody as a positive control reagent in an  
 CC immunoassay for the detection of anti HIV-1 group O antibodies. The  
 CC monoclonal antibodies are useful as positive control reagents in  
 CC immunoassays capable of detecting anti-HIV-1 group O antibodies. Such  
 CC immunoassays involve coupling a monoclonal antibody with HIV group-1  
 CC antigen and detecting the antigen-antibody complex. The monoclonal  
 CC antibodies of the invention would be used to ensure that the reagents  
 CC provided to detect HIV-1 group O antibody were performing properly. The  
 CC monoclonal antibodies may also be immobilised on a matrix and used  
 CC for affinity purification of specific HIV-1 group O-derived proteins from  
 CC cell cultures or biological tissues. The monoclonal antibodies can also  
 CC be used for generating chimeric antibodies for therapeutic use. Different  
 CC epitopes of HIV antigens can be used in combination in assay to diagnose,  
 CC evaluate, or prognosticate HIV disease condition. The monoclonal  
 CC antibodies are also useful for differentiating HIV-1 group O antigens  
 CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent  
 CC recombinant HIV-1 group O env antigens encoded by the synthetic genes  
 CC AAZ90280-290286. The recombinant HIV-1 env proteins contain various  
 CC deletions relative to the native HAM112 isolate env protein (AAY77376).  
 CC The recombinant HIV-1 group O antigens were purified and used to screen  
 CC hybridoma cultures  
 XX Sequence 215 AA;  
 SQ Query Match 91.5%; Score 161; DB 3; Length 215;  
 Best Local Similarity 84.4%; Pred. No. 4e-13;  
 Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALETLLNQQLLNSWGCGRGLVCYTSVRWNET 32  
 DB 117 ALETLLNQQLLNSWGCGRGLVCYTSVRWNET 148  
 RESULT 9  
 AAY09493  
 ID AAY09493 standard; protein; 245 AA.  
 XX AAY09493;  
 XX 17-OCT-2003 (revised)  
 DT 15-JUL-1999 (first entry)  
 XX HIV-1 Group O env polypeptide pGO-9PL.  
 XX HIV; human immunodeficiency virus; antigen; detection; antibody;  
 KW differentiation; Group O; env; immunogen; immunoassay.  
 XX Human immunodeficiency virus 1.  
 OS WO9909179-A2.  
 XX 25-FEB-1999.  
 PD 17-AUG-1998; 98WO-US017014.

XX 15-AUG-1997; 97US-00911824.  
 XX (ABBO ) ABBOTT LAB.  
 XX Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;  
 FI WPI; 1999-190167/16.  
 DR N-PSDB; AAX56074.  
 XX New isolated HIV-1 Group O env polypeptides - used for the detection of  
 PT anti-HIV antibodies and for the production of antibodies for use in  
 PT detection, purification and therapy.  
 XX Claim 16; Fig 7; 138pp; English.  
 XX The present invention describes (A) an isolated HIV-1 Group O env  
 CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env  
 CC polypeptide comprising an immunoreactive portion of a polypeptide as in  
 CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);  
 CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide  
 CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct  
 CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at  
 CC least one HIV-1 Group M env polypeptide; (5) an antigen construct  
 CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env  
 CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an  
 CC antigen construct comprising a first HIV-2 env polypeptide fused to a  
 CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as  
 CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a  
 CC host cell transformed by an expression vector as in (8); and (10) an  
 CC immunoassay kit for the detection of antibodies to HIV-1 comprising an  
 CC antigen construct as in (3)-(6). The antigen constructs can be used for  
 CC the detection of anti-HIV-1 antibodies in test samples. They can also be  
 CC used as immunogens to produce antibodies. The antibodies can be used to  
 CC purify HIV polypeptides, for therapy and for detection of HIV  
 CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)  
 XX Sequence 245 AA;  
 SQ Query Match 91.5%; Score 161; DB 2; Length 245;  
 Best Local Similarity 84.4%; Pred. No. 4.6e-13;  
 Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALETLLNQQLLNSWGCGRGLVCYTSVRWNET 32  
 DB 117 ALETLLNQQLLNSWGCGRGLVCYTSVRWNET 148  
 RESULT 10  
 AAY06977  
 ID AAY06977 standard; protein; 245 AA.  
 XX AAY06977;  
 XX 06-JUL-1999 (first entry)  
 DT Recombinant pGO-9PL protein.  
 DE HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;  
 KW antibody; assay.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX Key Location/Qualifiers  
 FH Protein 2..46  
 FT /note= "gp120 sequence"  
 FT Peptide 47..245  
 FT /note= "gp41 sequence"  
 XX WO9909410-A2.  
 XX 25-FEB-1999.  
 PD



PD	25-FEB-1999.
XX	
XX	17-AUG-1998; 98WO-US017014.
XX	
PR	15-AUG-1997; 97US-00911824.
XX	(ABBO ) ABBOTT LAB.
PA	
XX	Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
PI	WPI; 1999-190167/16:
XX	N-PSDB; AAX56132.
DR	
XX	New isolated HIV-1 Group O env polypeptides - used for the detection of
PT	anti-HIV antibodies and for the production of antibodies for use in
PT	detection, purification and therapy.
XX	
XX	Claim 57; Fig 17; 138pp; English.
XX	
CC	The present invention describes (A) an isolated HIV-1 Group O env
CC	polypeptide. Also described are: (1) an isolated HIV-1 Group O env
CC	polypeptide comprising an immunoreactive portion of a polypeptide as in
CC	(A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
CC	(3) an antigen construct comprising a first HIV-1 Group O env polypeptide
CC	fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
CC	comprising a fusion of at least one HIV-1 Group O env polypeptide with at
CC	least one HIV-1 Group M env polypeptide; (5) an antigen construct
CC	comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
CC	polypeptide, and at least one additional HIV-1 polypeptide; (6) an
CC	antigen construct comprising a first HIV-2 env polypeptide fused to a
CC	second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
CC	in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
CC	host cell transformed by an expression vector as in (8); and (10) an
CC	immunoassay kit for the detection of antibodies to HIV-1 comprising an
CC	antigen construct as in (3)-(6). The antigen constructs can be used for
CC	the detection of anti-HIV-1 antibodies in test samples. They can also be
CC	used as immunogens to produce antibodies. The antibodies can be used to
CC	purify HIV polypeptides, for therapy and for detection of HIV
CC	polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX	
SQ	Sequence 281 AA;
	Query Match 91.5%; Score 161; DB 2; Length 281;
	Best Local Similarity 84.4%; Pred. No. 5.3e-13;
	Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy	1 ALETLLNQQLLNSWGCRGLVCYTSVRWNKT 32
	:     :     :     :
Db	117 ALETLIQNQQLNLWGCKGRILICYTSVKWNKT 148
RESULT 13	
AAV09495	
ID	AAV09495 standard; protein; 373 AA.
XX	
AC	AAV09495;
DT	
DT	17-OCT-2003 (revised)
DT	15-JUL-1999 (first entry)
XX	
DE	HIV-1 Group O env polypeptide pGO-11PL.
XX	
KW	HIV; human immunodeficiency virus; antigen; detection; antibody;
KW	differentiation; Group O; env; immunogen; immunoassay.
OS	
OS	Human immunodeficiency virus 1.
XX	
PX	WO9909179-A2.
XX	
PD	25-FEB-1999.
XX	
PF	17-AUG-1998; 98WO-US017014.
XX	

```

PR 15-AUG-1997; 97US-00911824.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Hackett JR, Yanaguchi J, Golden AM, Brennan CA, Hickman RK;
XX
XX WPI; 1999-190167/16.
XX
XX DR N-PSDB; MAX56076.
XX
XX
XX New isolated HIV-1 Group O env polypeptides - used for the detection of
XX anti-HIV antibodies and for the production of antibodies for use in
XX detection, purification and therapy.
XX
XX Claim 15; Fig 9; 138pp; English.
XX
XX The present invention describes (A) an isolated HIV-1 Group O env
XX polypeptide. Also described are: (1) an isolated HIV-1 Group O env
XX polypeptide comprising an immunoreactive portion of a polypeptide as in
XX (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
XX (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
XX fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
XX comprising a fusion of at least one HIV-1 Group O env polypeptide with at
XX least one HIV-1 Group M env polypeptide; (5) an antigen construct
XX comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
XX polypeptide, and at least one additional HIV-1 polypeptide; (6) an
XX antigen construct comprising a first HIV-2 env polypeptide fused to a
XX second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
XX in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
XX host cell transformed by an expression vector as in (8); and (10) an
XX immunoassay kit for the detection of antibodies to HIV-1 comprising an
XX antigen construct as in (3)-(6). The antigen constructs can be used for
XX the detection of anti-HIV-1 antibodies in test samples. They can also be
XX used as immunogens to produce antibodies. The antibodies can be used to
XX purify HIV polypeptides, for therapy and for detection of HIV
XX polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 373 AA;
XX
XX Query Match 91.5%; Score 161; DB 2; Length 373;
XX Best Local Similarity 84.4%; Pred. No. 7.2e-13;
XX Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0
XX
XX Qy 1 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 32
XX |||||:|||||:|||||:|||||:|||||:
XX 117 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 148
XX
XX RESULT 14
XX AAY06979
XX ID AAY06979 standard; protein; 373 AA.
XX AC AAY06979;
XX
XX DT 06-JUL-1999 (first entry)
XX
XX XX Recombinant pGO-11PL protein.
XX
XX DE HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;
XX antibody; assay.
XX
XX OS Synthetic.
XX OS Human immunodeficiency virus 1.
XX
XX FH Key Location/Qualifiers
XX FT Protein 2..46
XX FT /note= "gp120 sequence"
XX FT Peptide 47..245
XX FT /note= "gp41 sequence"
XX
XX PN W09909410-A2.
XX
XX PD 25-FEB-1999.
XX

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 17.6134 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-11

Perfect score: 176

Sequence: 1 ALEFLNQQLLSWGCRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	95.5	40	3	US-08-894-699-39
2	168	95.5	40	3	US-08-444-410-39
3	161	91.5	215	2	US-08-912-129A-58
4	161	91.5	245	2	US-08-912-129A-48
5	161	91.5	373	2	US-08-912-129A-52
6	161	91.5	460	2	US-08-912-129A-60
7	161	91.5	490	2	US-08-912-129A-50
8	161	91.5	618	2	US-08-912-129A-54
9	161	91.5	873	2	US-08-912-129A-61
10	158	89.8	33	3	US-09-433-428D-25
11	158	89.8	40	3	US-08-894-699-36
12	158	89.8	40	3	US-08-894-699-68
13	158	89.8	40	3	US-08-444-410-36
14	158	89.8	40	3	US-09-444-410-68
15	158	89.8	116	4	US-09-462-917A-20
16	158	89.8	356	1	US-08-602-713-12
17	158	89.8	356	3	US-08-989-493-12
18	158	89.8	356	4	US-09-610-271-12
19	157	89.2	41	3	US-08-894-699-69
20	157	89.2	41	3	US-09-444-410-69
21	157	89.2	113	4	US-09-462-917A-2
22	157	89.2	117	4	US-09-462-917A-6
23	156	88.6	33	3	US-09-433-428D-6
24	156	88.6	33	3	US-09-433-428D-30
25	156	88.6	41	3	US-08-894-699-67
26	156	88.6	41	3	US-09-444-410-67
27	156	88.6	110	4	US-09-462-917A-14

Sequence 134, Appl  
Sequence 94, Appl  
Sequence 100, Appl  
Sequence 16, Appl  
Sequence 10, Appl  
Sequence 12, Appl  
Sequence 37, Appl  
Sequence 42, Appl  
Sequence 32, Appl  
Sequence 47, Appl  
Sequence 30, Appl  
Sequence 40, Appl  
Sequence 40, Appl  
Sequence 28, Appl  
Sequence 4, Appl  
Sequence 7, Appl  
Sequence 15, Appl

28 156 88.6 715 4 US-09-462-917A-134  
29 155 88.1 37 4 US-08-817-441-94  
30 155 88.1 104 4 US-08-817-441-100  
31 155 88.1 110 4 US-09-462-917A-16  
32 155 88.1 113 4 US-09-462-917A-10  
33 155 88.1 113 4 US-09-462-917A-12  
34 154 87.5 40 3 US-08-894-699-37  
35 154 87.5 40 3 US-08-894-699-42  
36 154 87.5 40 3 US-09-444-410-37  
37 154 87.5 40 3 US-09-444-410-42  
38 154 87.5 110 4 US-09-462-917A-30  
39 153 86.9 33 3 US-09-433-428D-10  
40 153 86.9 40 3 US-08-894-699-40  
41 153 86.9 40 3 US-09-444-410-40  
42 153 86.9 113 4 US-09-462-917A-28  
43 152 86.4 33 3 US-09-433-428D-4  
44 152 86.4 33 3 US-09-433-428D-7  
45 152 86.4 33 3 US-09-433-428D-15

#### ALIGNMENTS

RESULT 1  
US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1 FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
US-08-894-699-39

Query Match 95.5%; Score 168; DB 3; Length 40;  
Best Local Similarity 90.6%; Pred. No. 1.1e-17;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLSWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 9 ALETLLNQQLLSWGCRGLVCYTSVKWNET 40

## RESULT 2

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSET-ARAKA, IBITISSAM  
; APPLICANT: LY, THOI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/444,410  
; FILING DATE:  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/894,699  
; FILING DATE:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995

ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614

REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 95.5%; Score 168; DB 3; Length 40;  
Best Local Similarity 90.6%; Pred. No. 1.1e-17;  
Matches 29; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLSWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 9 ALETLLNQQLLSWGCRGLVCYTSVKWNET 40

## RESULT 3

US-08-912-129A-58  
; Sequence 58, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:

APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.

APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.

APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.

APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.

TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
; NUMBER OF SEQUENCES: 89

CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road  
; CITY: Abbott Park

STATE: IL  
; COUNTRY: USA

ZIP: 60064-3500  
; COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
; COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS (Windows 95)  
; SOFTWARE: Microsoft Word (ASCII format output)

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,129A

FILING DATE: 15-AUG-1997  
; CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:

FILING DATE:  
; ATTORNEY/AGENT INFORMATION:

NAME: Dancers, Andreas M.  
; REGISTRATION NUMBER: 32,652

REFERENCE/DOCKET NUMBER: 6109.US.01  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-937-9803  
; TELEFAX: 847-938-2623

TELEX:  
; INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 215 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-912-129A-58

Query Match 91.5%; Score 161; DB 2; Length 215;  
Best Local Similarity 84.4%; Pred. No. 8e-16;  
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLSWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 117 ALETLLNQQLLSWGCRGLVCYTSVKWNET 148

## RESULT 4

US-08-912-129A-48  
; Sequence 48, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:

APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.

APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.

APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.

APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.



```
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912.129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-48

Query Match          91.5%; Score 161; DB 2; Length 245;
Best Local Similarity 84.4%; Pred. No. 9.3e-16;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNSWCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLLNQQLLNLWGCKRGLCYTSVKWNET 148

RESULT 5
US-08-912-129A-52
; Sequence 52, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-48
```

```
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912.129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-52

Query Match          91.5%; Score 161; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 1.5e-15;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNSWCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLLNQQLLNLWGCKRGLCYTSVKWNET 148

RESULT 6
US-08-912-129A-60
; Sequence 60, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912.129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-912-129A-60

Query Match 91.5%; Score 161; DB 2; Length 460;
Best Local Similarity 84.4%; Pred. No. 1.9e-15;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32
Db 362 ALETLNQQLLNWGCGRGLVCYTSVKWNET 393

RESULT 7
US-08-912-129A-50
; Sequence 50, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-50

Query Match 91.5%; Score 161; DB 2; Length 490;
```

```
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-912-129A-60

Query Match 91.5%; Score 161; DB 2; Length 460;
Best Local Similarity 84.4%; Pred. No. 2.6e-15;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32
Db 362 ALETLNQQLLNWGCGRGLVCYTSVKWNET 393

RESULT 8
US-08-912-129A-54
; Sequence 54, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-54

Query Match 91.5%; Score 161; DB 2; Length 618;
Best Local Similarity 84.4%; Pred. No. 2.6e-15;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32
Db 362 ALETLNQQLLNWGCGRGLVCYTSVKWNET 393

RESULT 9
US-08-912-129A-61
; Sequence 61, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
```

; APPLICANT: VALLARI, ANADRUZELA S.  
 ; APPLICANT: HACKETT, JOHN JR.  
 ; APPLICANT: HICKMAN, ROBERT K.  
 ; APPLICANT: VARITEK, VINCENT A. JR.  
 ; APPLICANT: NECKLAWS, ELIZABETH A.  
 ; APPLICANT: GOLDEN, ALAN M.  
 ; APPLICANT: BRENNAN, CATHERINE A.  
 ; APPLICANT: DEVARE, SUSHIL G.  
 ; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
 ; NUMBER OF SEQUENCES: 89  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Abbott Laboratories  
 ; STREET: 100 Abbott Park Road  
 ; CITY: Abbott Park  
 ; STATE: IL  
 ; COUNTRY: USA  
 ; ZIP: 60064-3500  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: MS-DOS (Windows 95)  
 ; SOFTWARE: Microsoft Word (ASCII format output)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/912,129A  
 ; FILING DATE: 15-AUG-1997  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Danckers, Andreas M.  
 ; REGISTRATION NUMBER: 32,652  
 ; REFERENCE/DOCKET NUMBER: 6109.US.01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 847-937-9803  
 ; TELEFAX: 847-938-2623  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 61:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Protein  
 ; US-08-912-129A-61

Query Match 91.5%; Score 161; DB 2; Length 873;  
 Best Local Similarity 84.4%; Pred. No. 3.9e-15;  
 Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 32  
 Db 591 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 622

RESULT 10  
 ; Sequence 25, Application US/09433428D  
 ; Patent No. 6149910  
 ; GENERAL INFORMATION:  
 ; APPLICANT: De Leys, Robert J.  
 ; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
 ; FILE REFERENCE: CDS-207  
 ; CURRENT APPLICATION NUMBER: US/09/433,428D  
 ; CURRENT FILING DATE: 1999-11-04  
 ; NUMBER OF SEQ ID NOS: 70  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 25  
 ; LENGTH: 33  
 ; TYPE: PRT  
 ; ORGANISM: Human immunodeficiency virus type 1  
 ; US-09-433-428D-25

Query Match 89.8%; Score 158; DB 3; Length 33;  
 Best Local Similarity 90.0%; Pred. No. 2.7e-16;  
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 30  
 Db 4 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 33

RESULT 11  
 ; US-08-894-699-36  
 ; Sequence 36, Application US/08894699  
 ; Patent No. 6030769  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SIMON, FRANCOIS  
 ; APPLICANT: SARAGOSTI, SENTOB  
 ; APPLICANT: LOUSERT-AJAKA, IBITISSAM  
 ; APPLICANT: LY, THOAI-DUONG  
 ; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
 ; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
 ; NUMBER OF SEQUENCES: 81  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
 ; CITY: FLOOR  
 ; STATE: ARLINGTON  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/894,699  
 ; FILING DATE: 01-DEC-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/FR96/00294  
 ; FILING DATE: 26-FEB-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 95/02236  
 ; FILING DATE: 27-FEB-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,614  
 ; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 36:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 40 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-894-699-36

Query Match 89.8%; Score 158; DB 3; Length 40;  
 Best Local Similarity 84.4%; Pred. No. 3.3e-16;  
 Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 32  
 Db 9 ALETLLNQQLLNSWGCRGLVCYTSVRWNET 40

RESULT 12

```
US-08-894-699-68
; Sequence 68, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-OPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-699-68

Query Match 89.8%; Score 158; DB 3; Length 40;
Best Local Similarity 87.5%; Pred. No. 3.3e-16;
Matches 28; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWCGRGLVCYTSVRWNET 32
Db 9 ALETLNQQLNSWCGKGRVCYTSVKNWT 40

RESULT 13
US-09-444-410-36
; Sequence 36, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-OPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-444-410-36

Query Match 89.8%; Score 158; DB 3; Length 40;
Best Local Similarity 84.4%; Pred. No. 3.3e-16;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWCGRGLVCYTSVRWNET 32
Db 9 ALETLNQQLNSWCGKGRVCYTSVKNWT 40

RESULT 14
US-09-444-410-68
; Sequence 68, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-OPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-444-410-36
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 89.8%; Score 158; DB 3; Length 40;  
Best Local Similarity 87.5%; Pred. NO. 3.3e-16;  
Matches 28; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWGCKGRVLCVTSVRWNET 32  
Db 9 ALETLLNQQLNSWGCKGRVLCVTSVRWNET 40

## RESULT 15

US-09-462-917A-20  
; Sequence 20, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Human  
US-09-462-917A-20

Query Match 89.8%; Score 158; DB 4; Length 116;  
Best Local Similarity 84.4%; Pred. NO. 1.1e-15;  
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWGCKGRVLCVTSVRWNET 32  
Db 29 ALETLLNQQLNSWGCKGRVLCVTSVRWNET 60

Search completed: May 7, 2004, 17:53:26  
Job time : 17.6134 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 44.7731 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-11

Perfect score: 176  
Sequence: 1 ALETLNQQLLNSWGCGRVLCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 27566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161	91.5	215	8	US-08-911-824-58
2	161	91.5	245	8	US-08-911-824-48
3	161	91.5	281	8	US-08-911-824-120
4	161	91.5	373	8	US-08-911-824-52
5	161	91.5	460	8	US-08-911-824-60
6	161	91.5	488	8	US-08-911-824-95
7	161	91.5	490	8	US-08-911-824-50
8	161	91.5	526	8	US-08-911-824-97
9	161	91.5	618	8	US-08-911-824-54
10	161	91.5	706	8	US-08-911-824-93
11	161	91.5	736	8	US-08-911-824-91
12	161	91.5	873	8	US-08-911-824-61
13	158	89.8	116	14	US-10-320-786-20
14	158	89.8	356	14	US-10-357-400-12
15	157	89.2	113	14	US-10-320-786-2

16	157	89.2	117	14	US-10-320-786-6	Sequence 6, Appli
17	156	88.6	110	14	US-10-320-786-14	Sequence 14, Appl
18	156	88.6	715	14	US-10-320-786-134	Sequence 134, App
19	155	88.1	37	14	US-10-026-741-94	Sequence 94, Appl
20	155	88.1	104	14	US-10-026-741-100	Sequence 100, App
21	155	88.1	110	14	US-10-320-786-16	Sequence 16, Appl
22	155	88.1	113	14	US-10-320-786-10	Sequence 10, Appl
23	155	88.1	113	14	US-10-320-786-12	Sequence 12, Appl
24	154	87.5	110	14	US-10-320-786-30	Sequence 30, Appl
25	153	86.9	113	14	US-10-320-786-28	Sequence 28, Appl
26	152	86.4	113	14	US-10-320-786-18	Sequence 18, Appl
27	152	86.4	116	14	US-10-320-786-22	Sequence 22, Appl
28	151	85.8	37	14	US-10-026-741-86	Sequence 86, Appl
29	151	85.8	200	9	US-09-854-816-104	Sequence 104, App
30	151	85.8	862	14	US-10-369-294-15	Sequence 15, Appl
31	150	85.2	115	14	US-10-320-786-24	Sequence 24, Appl
32	148	84.1	113	14	US-10-320-786-40	Sequence 40, Appl
33	147	83.5	35	14	US-10-026-741-101	Sequence 101, App
34	147	83.5	37	14	US-10-026-741-7	Sequence 7, Appli
35	147	83.5	37	14	US-10-026-741-88	Sequence 88, Appl
36	147	83.5	37	14	US-10-026-741-90	Sequence 90, Appl
37	147	83.5	110	14	US-10-320-786-36	Sequence 36, Appl
38	147	83.5	351	14	US-10-026-741-47	Sequence 47, Appl
39	147	83.5	877	14	US-10-026-741-102	Sequence 102, App
40	142	80.7	35	9	US-09-886-156-62	Sequence 62, Appl
41	142	80.7	35	9	US-09-886-150-62	Sequence 62, Appl
42	142	80.7	35	10	US-09-886-149-62	Sequence 62, Appl
43	142	80.7	35	10	US-09-886-159-62	Sequence 62, Appl
44	142	80.7	35	14	US-10-326-090-62	Sequence 62, Appl
45	142	80.7	146	13	US-10-000-321-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 91.5%; Score 161; DB 8; Length 215;

Best Local Similarity 84.4%; Pred No. 5.4e-14;  
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNSWGCGRVLCYTSVRWNET 32

Db 117 ALETLNQQLLNSWGCGRVLCYTSVRWNET 148

RESULT 2

US-08-911-824-48





```

; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-14PL
US-08-911-824-95

Query Match          91.5%; Score 161; DB 8; Length 488;
Best Local Similarity 84.4%; Pred. No. 1.2e-13;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLNLWGCKGRLCYTSVKWNET 148

RESULT 7
US-08-911-824-50
; Sequence 50, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-50

Query Match          91.5%; Score 161; DB 8; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.2e-13;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 362 ALETLNQQLNLWGCKGRLCYTSVKWNET 393

RESULT 8
US-08-911-824-97
; Sequence 97, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS
US-08-911-824-97

Query Match          91.5%; Score 161; DB 8; Length 526;
Best Local Similarity 84.4%; Pred. No. 1.3e-13;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 362 ALETLNQQLNLWGCKGRLCYTSVKWNET 393

RESULT 9
US-08-911-824-54
; Sequence 54, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

Query Match          91.5%; Score 161; DB 8; Length 618;
Best Local Similarity 84.4%; Pred. No. 1.5e-13;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 362 ALETLNQQLNLWGCKGRLCYTSVKWNET 393

RESULT 10
US-08-911-824-93
; Sequence 93, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-93

Query Match          91.5%; Score 161; DB 8; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.2e-13;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 362 ALETLNQQLNLWGCKGRLCYTSVKWNET 393
```



NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/357,400  
FILING DATE: 04-Feb-2003  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,271  
FILING DATE: 06-Jul-2000  
APPLICATION NUMBER: 08/602,713  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No US20030147917Alman D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LEDER 203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-357-400-12

Query Match 89.8%; Score 158; DB 14; Length 356;  
Best Local Similarity 81.2%; Pred. No. 2.3e-13;  
Matches 26; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLQNQQLLNSWGCGRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 312 ALETLMQNQQLLNLWGCKGLICYTSVKWNET 343

RESULT 15  
US-10-320-786-2  
Sequence 2, Application US/10320786  
Publication No. US20030180759A1  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
APPLICANT: Vanden Haesevelde, Marlen  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 2  
LENGTH: 113  
TYPE: PRT  
ORGANISM: Human  
US-10-320-786-2

Query Match 89.2%; Score 157; DB 14; Length 113;  
Best Local Similarity 81.2%; Pred. No. 1e-13;  
Matches 26; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ALETLLQNQQLLNSWGCGRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 29 ALETLLQNQQLLNLWGCKGRIVCYTSVKWNT 60

Search completed: May 7, 2004, 18:29:21  
Job time : 44.7731 secs

THIS PAGE BLANK (USP10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 13.4454 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-11  
Perfect score: 176  
Sequence: 1 ALETLNQQLLSWGCRGLVCYTSVRWNET 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155	88.1	104	2 S52930	GP41 ENV protein -
2	151	85.8	863	2 A53034	gag polyprotein -
3	147	83.5	877	2 S49197	env polyprotein p
4	110	62.5	854	1 VCLJSH	env polyprotein pr
5	105	59.7	357	2 S21990	env polyprotein g
6	105	59.7	358	2 S22002	env polyprotein g
7	105	59.7	454	2 B41621	env polyprotein D
8	104	59.1	358	2 S22000	env polyprotein g
9	104	59.1	358	2 S70417	env polyprotein g
10	103	58.5	443	2 C41621	env polyprotein p
11	103	58.5	853	2 S54384	env polyprotein pr
12	103	58.5	855	1 VCLJZR	env polyprotein pr
13	102	58.0	357	2 S22006	env polyprotein g
14	102	58.0	357	2 S21994	env polyprotein g
15	102	58.0	357	2 S22004	env polyprotein g
16	102	58.0	357	2 S21996	env polyprotein g
17	102	58.0	357	2 S21992	env polyprotein g
18	102	58.0	358	2 S21998	env polyprotein g
19	102	58.0	445	2 A41621	env polyprotein M
20	102	58.0	843	1 H44001	env polyprotein pr
21	102	58.0	852	1 VCLJBR	env polyprotein -
22	102	58.0	852	2 T12016	env polyprotein -
23	102	58.0	854	2 S13288	env protein - huma
24	102	58.0	855	1 VCLJJA2	env polyprotein pr
25	102	58.0	856	1 VCLJH3	env polyprotein pr
26	102	58.0	856	1 VCLJVL	env polyprotein pr
27	102	58.0	856	1 VCLJ3W	env polyprotein pr
28	102	58.0	859	1 VCLJWN	env polyprotein pr
29	102	58.0	861	1 VCLJLV	env polyprotein pr

30	102	58.0	861	1 VCLJSC	env polyprotein pr
31	102	58.0	868	1 VCLJH4	env polyprotein -
32	101.5	57.7	855	2 A45713	Env transmembrane
33	100	56.8	856	1 A44963	env polyprotein pr
34	99	56.2	846	1 VCLJND	env polyprotein pr
35	99	56.2	847	2 T09448	env polyprotein pr
36	99	56.2	847	2 S13289	env polyprotein - huma
37	96	54.5	729	1 VCLJJK	env polyprotein pr
38	96	54.5	861	1 VCLJKB	env polyprotein pr
39	95.5	54.3	859	1 VCLJST	env polyprotein pr
40	95.5	54.3	859	2 S24571	env protein - huma
41	95.5	54.3	885	2 S04322	env polyprotein -
42	95.5	54.3	886	2 T11555	env protein - simi
43	95	54.0	851	2 S33985	env polyprotein -
44	94	53.4	858	1 VCLJG2	env polyprotein pr
45	94	53.4	859	2 T01672	envelope polyprote

ALIGNMENTS

RESULT 1

S52930

GP41 ENV protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C:Accession: S52930

R:Cohen, J.H.M.; Guebard, D.; Philibert, F.; Charnaret, S.; Tabary, T.; Montagnier, L.; de

submitted to the EMBL Data Library, January 1995

A:Description: A novel HIV1-O strain illustrates the diversity of the O group.

A:Reference number: S52929

A:Accession: S52930

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-104 <COH>

A:Cross-references: EMBL:X84328; NID:G95526; PIDN:CAA59066.1; PID:G695527

C:Superfamily: type E retrovirus env polyprotein

Query Match 88.1%; Score 155; DB 2; Length 104;

Best Local Similarity 81.2%; Pred. No. 1.2e-14;

Matches 26; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLSWGCRGLVCYTSVRWNET 32

Db 35 ALETLNQQLLSWGCRGLVCYTSVRWNET 66

RESULT 2

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborcht, B.; van der Groen, B.

J. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African hu

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A>Status: preliminary

A:Molecule type: Genomic RNA

A:Residues: 1-863 <VAN>

A:Cross-references: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 85.8%; Score 151; DB 2; Length 863;

Best Local Similarity 81.2%; Pred. No. 3.5e-13;

Matches 26; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLSWGCRGLVCYTSVRWNET 32

Db 584 ALETLNQQLLSWGCRGLVCYTSVRWNET 615

```
RESULT 3
S49197
envelope protein precursor - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 26-Aug-1999
C;Accession: S49197
R;Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy,
submitted to the EMBL Data Library, July 1994
A;Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: def1
A;Reference number: S49197
A;Accession: S49197
A;Molecule type: DNA
A;Residues: 1-877 <HUE>
A;Cross-references: EMBL:X80020; NID:g510516; PIDN:CAA56323.1; PID:g510517
A;Experimental source: isolate VAU
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-535/Product: coat protein gp120 #status predicted <CP1>
F;536-877/Product: coat protein gp41 #status predicted <CP2>
F;698-716/Domain: transmembrane #status predicted <TMN>
F;59,88,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404,
Query Match 83.5%; Score 147; DB 2; Length 877;
Best Local Similarity 75.0%; Pred. No. 1.3e-12;
Matches 24; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 ALETLNQQLLNWSGCGRLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 597 ALETFIQQLLNWLGCGNRILCYTSVRWNET 628

RESULT 4
VCLJ51
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: simian immunodeficiency virus SIVcpz
A;Note: host Pan troglodytes (chimpanzee)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: S09990
R;Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A;Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A;Reference number: S09983; MUID:90259077; PMID:2188136
A;Accession: S09990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-854 <HUE>
A;Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36407.1; PID:g58874
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-500/Product: coat protein gp120 #status predicted <CP1>
F;501-854/Product: coat protein gp41 #status predicted <CP2>
F;501-517/Domain: transmembrane #status predicted <TM1>
F;675-693/Domain: transmembrane #status predicted <TM2>
F;805-821/Domain: transmembrane #status predicted <TM3>
F;134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,45
Query Match 62.5%; Score 110; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 2e-07;
Matches 18; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 ALETLNQQLLNWSGCGRLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:
Db 572 AVERYLDQQLGLWGCGRKAVCYTTPWNNS 603

RESULT 5
env polyprotein D - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
```

```
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 20
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S21990; S70423
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
A;Reference number: S21990
A;Accession: S21990
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebro
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70423
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332,'X',334-357 <STE2>
A;Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C;Superfamily: type E retrovirus env polyprotein
Query Match 59.7%; Score 105; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 ALETLNQQLLNWSGCGRLVCYTSVRWN 30
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 83 AVERYLKDQQLGLWGCGRILCTTAVPN 112

RESULT 6
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 3L
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S22002; S70418
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
A;Reference number: S21990
A;Accession: S22002
A;Molecule type: DNA
A;Residues: 1-358 <STE1>
A;Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebro
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70418
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333,'X',335-358 <STE2>
A;Cross-references: EMBL:X61352; NID:g60186
C;Superfamily: type E retrovirus env polyprotein
Query Match 59.7%; Score 105; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 ALETLNQQLLNWSGCGRLVCYTSVRWN 30
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 84 AVERYLKDQQLGLWGCGRILCTTAVPN 113

RESULT 7
B41621
env polyprotein D - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
```

```
A;Cross-references: EMBL:X61351; NID:g60184; PIDN:CAA43614.1.; PID:g60185  
C;Superfamily: type B retrovirus env polyprotein
```

Query Match	59.1%; Score 104;
Best Local Similarity	60.0%; Pred.No. 6.le-07;
Matches	Conservative    4; Mismatches    8; Indels         0; Gaps              0;

QY          1 ALETLNQQLNSWGCGRGLVCYTTSVRWN 30  
| : | : ||||| ||| ||||| ||  
DB        84 AVERYLKDQLLGWGSGRLICTTVPWN 113

RESULT 10  
CA1621  
env polypeptide P - human immunodeficiency virus type 1 (fragment)  
N;Alternate names: coat polyprotein  
N;Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999  
C;Accession: C41621  
R;Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991  
A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence d  
A;Reference number: CA1621; PMID:92107924; PMID:1763038  
A;Accession: C41621  
A;Molecule type: DNA  
A;Residues: 1-443 <BUR>  
A;Cross-references: GB:M7230; NID:g328631; PIDN:AAB03792.1; PID:g555015  
A;Note: this virus was isolated from the mother's sexual partner  
C;Genetics:  
A;Gene: env  
C;Superfamily: type B retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency;  
F;1-25/Product: coat protein gp120 (fragment) #status predicted <GP1>  
F;252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>  
F;424-443/Domain: transmembrane #status predicted <TMN>  
F;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding

Query Match	58.5%; Score 103;
Best Local Similarity	53.1%; Pred.No. 1e-06;
Matches	Conservative    7; Mismatches    8; Indels         0; Gaps              0;

QY          1 ALETLNQQLNSWGCGRGLVCYTTSVRWNET 32  
| : | : ||||| ||||| ||||| ||  
DB        322 AVERYLDQLLGWCSSKLICTTVPNWS 353

RESULT 11  
S54384  
envelope polyprotein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Aug-1999  
C;Accession: S54384  
R;Theodore, T.; Buckler-White, A.J.  
submitted to the EMBL Data Library, July 1989  
A;Reference number: S54377  
A;Accession: S54384  
A>Status: preliminary  
A;Molecule type: genomic RNA  
A;Residues: 1-853 <THE>  
A;Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: polyprotein

Query Match	58.5%; Score 103;
Best Local Similarity	53.1%; Pred.No. 1.9e-06;
Matches	Conservative    7; Mismatches    8; Indels         0; Gaps              0;

QY          1 ALETLNQQLNSWGCGRGLVCYTTSVRWNET 32  
| : | : ||||| ||||| ||||| ||  
DB        579 AVERYLKDQLLGWCSSKLICTTTVFNWS 610

```
RESULT 12
VCUJZR
env polyprotein precursor - human immunodeficiency virus Zr-6
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus Zr-6
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: D26192
R;Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu
Gene 52, 71-82, 1987
A;Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti
A;Reference number: A26192; MUID:87248097; PMID:3036660
A;Accession: D26192
A;Molecule type: DNA
A;Residues: 1-855 <SRI>
A;Cross-references: GB:X03458; GB:M16322; NID:G329398; PIDN:AAA45380.1; PID:G329403
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-855/Product: env polyprotein #status predicted <MAT>
F;501-855/Product: exterior membrane glycoprotein #status predicted <EXT>
F;87-129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,404
Query Match 58.5%; Score 103; DB 1; Length 855;
Best Local Similarity 53.1%; Pred. No. 2e-06;
Matches 17; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 ALETLNQQLNSWGCGRGLVCYTSVRWN 32
Db 581 AVERYLKDQQLLGWCGSKLICITTTVPWN 612

RESULT 13
S22006
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S70420; S22006
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <ST2>
A;Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191
A;Experimental source: patient L
A;Note: submitted to the EMBL Data Library, July 1991
C;Superfamily: type E retrovirus env polyprotein

Query Match 58.0%; Score 102; DB 2; Length 357;
Best Local Similarity 56.7%; Pred. No. 1.2e-06;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 1 ALETLNQQLNSWGCGRGLVCYTSVRWN 30
Db 83 AVERYLKDQQLLGWCGSKLICITTTVPWN 112

RESULT 14
S21994
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 27B
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S21994; S70421
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
```

```
A;Reference number: S21990
A;Accession: S21994
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140,'X',142-312,'X',314-357 <STE2>
A;Cross-references: EMBL:X61355; NID:g60179
C;Superfamily: type E retrovirus env polyprotein

Query Match 58.0%; Score 102; DB 2; Length 357;
Best Local Similarity 56.7%; Pred. No. 1.2e-06;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 1 ALETLNQQLNSWGCGRGLVCYTSVRWN 30
Db 83 AVERYLKDQQLLGWCGSKLICITTTVPWN 112

RESULT 15
S22004
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 4B
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S22004; S70419
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
A;Reference number: S21990
A;Accession: S22004
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-292,'X',294-357 <STE2>
A;Cross-references: EMBL:X61353; NID:g60188
C;Superfamily: type E retrovirus env polyprotein

Query Match 58.0%; Score 102; DB 2; Length 357;
Best Local Similarity 56.7%; Pred. No. 1.2e-06;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 1 ALETLNQQLNSWGCGRGLVCYTSVRWN 30
Db 83 AVERYLKDQQLLGWCGSKLICITTTVPWN 112

Search completed: May 7, 2004, 17:51:02
Job time : 13.4454 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 8.06723 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-11

Perfect score: 176

Sequence: 1 ALETLLQQLNSWGCRGLVCTVSRWNET 32

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	62.5	854	1 ENV_SIVCZ	P17281 chimpanzee
2	103	58.5	853	1 ENV_HV122	P12487 human immun
3	103	58.5	855	1 ENV_HV126	P04580 human immun
4	102	58.0	843	1 ENV_HV1Y2	P35961 human immun
5	102	58.0	847	1 ENV_HV1S1	P19550 human immun
6	102	58.0	847	1 ENV_HV1W2	P05880 human immun
7	102	58.0	851	1 ENV_HV1B8	P04582 human immun
8	102	58.0	852	1 ENV_HV1B9	P12488 human immun
9	102	58.0	852	1 ENV_HV1S3	P19549 human immun
10	102	58.0	853	1 ENV_HV1MF	P19551 human immun
11	102	58.0	855	1 ENV_HV1A2	P03378 human immun
12	102	58.0	855	1 ENV_HV1OY	P20888 human immun
13	102	58.0	856	1 ENV_HV1B1	P03375 human immun
14	102	58.0	856	1 ENV_HV1H2	P04578 human immun
15	102	58.0	856	1 ENV_HV1LW	Q70626 human immun
16	102	58.0	856	1 ENV_HV1MN	P05877 human immun
17	102	58.0	856	1 ENV_HV1PV	P03376 human immun
18	102	58.0	856	1 ENV_HV1SC	P05878 human immun
19	102	58.0	856	1 ENV_HV1W1	P31872 human immun
20	102	58.0	861	1 ENV_HV1B9	P03377 human immun
21	102	58.0	865	1 ENV_HV1RH	P04579 human immun
22	102	58.0	867	1 ENV_HV1J3	P12489 human immun
23	102	58.0	868	1 ENV_HV1C4	P05879 human immun
24	100	56.8	848	1 ENV_HV1JR	P20871 human immun
25	100	56.8	856	1 ENV_HV1H3	P04624 human immun
26	100	56.8	856	1 ENV_HV1ZH	P05881 human immun
27	99	56.2	846	1 ENV_HV1ND	P18799 human immun
28	98	55.7	863	1 ENV_HV1Z8	P05882 human immun
29	96	54.5	853	1 ENV_HV1EL	P04581 human immun
30	96	54.5	861	1 ENV_HV1KB	P31819 human immun
31	95.5	54.3	856	1 ENV_HV2NZ	P05883 human immun
32	95.5	54.3	859	1 ENV_HV2D2	P15831 human immun
33	95.5	54.3	885	1 ENV_SIVS4	P12492 simian immun

#### ALIGNMENTS

RESULT 1

ENV_SIVCZ	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
ID	ENV_SIVCZ			
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."			
RL	Nature 345:356-359(1990).			
CC	-!- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X52154; CAA36407.1; -.			
DR	PIR; S09990; VCLJ51.			
DR	HIV; X52154; ENVSCP2.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;			
KW	Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .) (POTENTIAL).

34	95	54.0	857	1	ENV_HV2KR	Q74126 human immun
35	94	53.4	858	1	ENV_HV2RO	P04577 human immun
36	94	53.4	859	1	ENV_HV1MA	P04583 human immun
37	93	52.8	865	1	ENV_SIVAT	P05886 simian immun
38	92.5	52.6	712	1	ENV_HV2S2	P32536 human immun
39	92.5	52.6	859	1	ENV_HV2ST	P20872 human immun
40	91.5	52.0	859	1	ENV_HV2CA	P24105 human immun
41	91	51.7	854	1	ENV_SIVAI	Q02837 simian immun
42	90.5	51.4	846	1	ENV_HV2SB	P12449 human immun
43	90	51.1	821	1	ENV_SIVGB	P22380 simian immun
44	90	51.1	851	1	ENV_HV2D1	P17755 human immun
45	90	51.1	851	1	ENV_HV2G1	P18040 human immun





Query Match 58.0%; Score 102; DB 1; Length 843;  
Best Local Similarity 56.7%; Pred. No. 2e-07;  
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALETLQNQLNSGCRGLVCYTSPVWN 30  
| | : : : : : ||||| : : : : : ||  
569 AVERYLDQQLGIWGCSKGLICTTAVPWN 598

Db

RESULT 5  
ENV\_HV1S1 STANDARD; PRT; 847 AA.  
AC PL9550;  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11691;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90347835; PubMed=2384920;  
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;  
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";  
RJ J. Virol. 64:4390-4398 (1990).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements/send an email to license@ebi.ac.uk).

EMBL; M65024; AAA45072.1; --  
PDB; LOBE; 15-MAY-97  
DR HIV; M38428; ENV\$SF162.  
DR InterPro; IPR000328; Env.GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.

KW SIGNAL 1 29  
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 203 BY SIMILARITY.  
FT DISULFID 125 194 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 216 245 BY SIMILARITY.  
FT DISULFID 226 237 BY SIMILARITY.  
FT DISULFID 294 328 BY SIMILARITY.  
FT DISULFID 374 435 BY SIMILARITY.  
FT DISULFID 381 408 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 58.0%; Score 102; DB 1; Length 847;  
Best Local Similarity 56.7%; Pred. No. 2e-07;  
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALETLQNQLNSGCRGLVCYTSPVWN 30  
| | : : : : : ||||| : : : : : ||  
573 AVERYLDQQLGIWGCSKGLICTTAVPWN 602

Db

RESULT 6  
ENV\_HV1W2 STANDARD; PRT; 847 AA.  
AC P05880;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11705;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86235450; PubMed=3012778;  
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D., Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS.";  
RL Science 232:1548-1553 (1986).  
CC -! MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements/send an email to license@ebi.ac.uk).

EMBL; M12507; AAB12990.1; --  
HIV; M12507; ENV\$WMJ2.  
DR InterPro; IPR000328; Env.GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.

KW SIGNAL 1 29  
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 202 BY SIMILARITY.  
FT DISULFID 125 193 BY SIMILARITY.  
FT DISULFID 130 152 BY SIMILARITY.  
FT DISULFID 215 244 BY SIMILARITY.  
FT DISULFID 225 236 BY SIMILARITY.  
FT DISULFID 293 326 BY SIMILARITY.

FT DISULFID 372 435 BY SIMILARITY. (POTENTIAL).  
FT DISULFID 379 408 BY SIMILARITY. (POTENTIAL).  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 847 AA; 96466 MW; CDLE33D73AASBCAE CRC64;  
Query Match 58.0%; Score 102; DB 1; Length 847;  
Best Local Similarity 56.7%; Pred. No. 2e-07;  
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
Oy 1 ALETLLNQQLNSWGRGLVCYTSVRWN 30  
Db 573 AVERYLKQQLLGWCGSKLICTTTPWN 602  
RESULT 7  
ID ENV HV1B8 STANDARD; PRT; 851 AA.  
AC P04582;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (BHS isolate) (HIV-1).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11684;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8511123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Pettey S.R. Jr., Pearson M.L., Lautenberger J.A., Papas T.S., Graybe J., Chang N.T., Gallo R.C., Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; K02011; AAA44661.1; -.  
DR PDB; 1DDH; 13-JAN-99.  
DR PDB; 1HHG; 31-OCT-93.  
DR PDB; 1QO3; 02-JAN-00.  
DR PDB; 1S2T; 24-DEC-97.  
DR HIV; K02011; ENV\$B8.  
DR GlycosuitedB; P04582; -.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
KW 3D-structure.  
FT CHAIN 1 30  
FT CHAIN 31 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT DISULFID 507 851 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 205 BY SIMILARITY.  
FT DISULFID 126 196 BY SIMILARITY.  
FT DISULFID 131 157 BY SIMILARITY.  
FT DISULFID 218 247 BY SIMILARITY.  
FT DISULFID 228 239 BY SIMILARITY.  
FT DISULFID 296 331 BY SIMILARITY.  
FT DISULFID 378 440 BY SIMILARITY.  
FT DISULFID 385 413 BY SIMILARITY.  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;  
Query Match 58.0%; Score 102; DB 1; Length 851;  
Best Local Similarity 56.7%; Pred. No. 2e-07;  
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
Oy 1 ALETLLNQQLNSWGRGLVCYTSVRWN 30  
Db 577 AVERYLKQQLLGWCGSKLICTTTPWN 606  
RESULT 8  
ID ENV HV1B8 STANDARD; PRT; 852 AA.  
AC P12488;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane



```

FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 852 AA; 96668 MW; EE7BBF8D23C9910D CRC64;

Query Match 58.0%; Score 102; DB 1; Length 852;
Best Local Similarity 56.7%; Pred. No. 2e-07;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWCGRLVCVTSVRWN 30
Db 578 AVERYLRDQQLGIGWCGSLCTTAVPWN 607

RESULT 10
ENV_HV1MFP
ID ENV_HV1MFP STANDARD; PRT; 853 AA.
AC P19551;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonic C., Mann A.M., Meier C.,
RA Wasia A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
cytotoxicity."
RT J. Virol. 64:3792-3803 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33943; AAA44850.1; --
CC PDB; 1A1K; 16-JUN-97.
CC HIV; M33943; ENV5MFA.
CC InterPro; IPR000328; Env GP41.
CC Dr InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30

```

```

FT CHAIN 31 31 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 510 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 74 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 337B993B6F22ABA CRC64;

Query Match 58.0%; Score 102; DB 1; Length 853;
Best Local Similarity 56.7%; Pred. No. 2e-07;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWCGRLVCVTSVRWN 30
Db 580 AVERYLRDQQLGIGWCGSLCTTAVPWN 609

RESULT 11
ENV_HV1A2
ID ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
(RARV-2).";
RT Science 227:484-492 (1985).

```



```
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; K02007; AAB59882.1; -.
DR PIR; A03976; VCLJA2.
DR HIV; K02007; ENV5SF2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 208 BY SIMILARITY.
FT DISULFID 125 199 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 221 250 BY SIMILARITY.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 299 333 BY SIMILARITY.
FT DISULFID 380 442 BY SIMILARITY.
FT DISULFID 387 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL. .)
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;
Query Match 58.0%; Score 102; DB 1; Length 855;
Best Local Similarity 56.7%; Pred. No. 2e-07;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
Qy 1 ALETLNQQLNSWGCRGLVCVTSVRWN 30
Db 581 AVERYRDQQLIGWGSGKGLICTAVPWN 610
RESULT 12
```

```
ENV_HVIOY STANDARD; PRT; 855 AA.
ID ENV_HVIOY
AC F20888;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet P., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -!- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A
HEALTHY GABONESE INDIVIDUAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M26727; AAB83397.1; -.
DR HIV; M26727; ENV5OYI.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 210 BY SIMILARITY.
FT DISULFID 125 201 BY SIMILARITY.
FT DISULFID 130 162 BY SIMILARITY.
FT DISULFID 223 252 BY SIMILARITY.
FT DISULFID 233 244 BY SIMILARITY.
FT DISULFID 301 335 BY SIMILARITY.
FT DISULFID 381 442 BY SIMILARITY.
FT DISULFID 388 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL. .)
```



```

FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 58.0%; Score 102; DB 1; Length 855;
Best Local Similarity 56.7%; Pred. No. 2e-07;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWCGRGLVCYTSVRWN 30
Db 581 AVERYLKQQLLGWCGSKLICTTTPWN 610

RESULT 13
ENV_HV1B1
ID ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1996 (Rel. 01, Last sequence update)
DT 21-JUL-1996 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Ratafski J.A., Whitehorn E.A.,
RA Baumesler K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284 (1985).
RN [2]
RN DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
potential glycosylation sites of the type 1 recombinant human
immunodeficiency virus envelope glycoprotein (gp120) expressed in
Chinese hamster ovary cells.";
RL J. Biol. Chem. 265:10373-10382 (1990).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15654; AAA44205.1;
DR PIR; A03973; VCLJH3.
DR HIV; M15654; ENV03BH102.
DR InterPro; IPR00328; Env_GP41.
DR InterPro; IPR00077; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157

```

```

FT DISULFID 218 247 N-LINKED (GLCNAC. . .)
FT DISULFID 228 239 N-LINKED (GLCNAC. . .)
FT DISULFID 296 331 N-LINKED (GLCNAC. . .)
FT DISULFID 378 445 N-LINKED (GLCNAC. . .)
FT DISULFID 385 418 N-LINKED (GLCNAC. . .)
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .)
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .)
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .)
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .)
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .)
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .)
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .)
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

Query Match 58.0%; Score 102; DB 1; Length 856;
Best Local Similarity 56.7%; Pred. No. 2e-07;
Matches 17; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNSWCGRGLVCYTSVRWN 30
Db 582 AVERYLKQQLLGWCGSKLICTTTPWN 611

RESULT 14
ENV_HV1B2
ID ENV_HV1B2 STANDARD; PRT; 856 AA.
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69 (1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

```





100% PAGE BLANK (SFTO)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	168	95.5	216	15	Q9IEC5	Q9IEC5 human	immun
2	165	93.8	116	15	Q7ZJN9	Q7ZJN9 human	immun
3	163	92.6	219	15	Q9IBB6	Q9IBB6 human	immun
4	161	91.5	124	15	Q9IHU7	Q9IHU7 human	immun
5	161	91.5	126	15	Q9IHV1	Q9IHV1 human	immun
6	161	91.5	130	15	Q9IHU9	Q9IHU9 human	immun
7	161	91.5	135	15	Q9DQL9	Q9DQL9 human	immun
8	161	91.5	172	15	Q9IEB3	Q9IEB3 human	immun
9	161	91.5	234	15	Q9IEC2	Q9IEC2 human	immun
10	161	91.5	872	15	Q8Q7H0	Q8Q7H0 human	immun
11	161	91.5	882	15	Q8Q7F9	Q8Q7F9 human	immun
12	161	91.5	887	15	Q8Q7G9	Q8Q7G9 human	immun
13	160	90.9	240	15	Q9IE32	Q9IE32 human	immun
14	159	90.3	125	15	Q9IHU8	Q9IHU8 human	immun
15	159	90.3	216	15	Q9IEA5	Q9IEA5 human	immun
16	159	90.3	242	15	Q9IE31	Q9IE31 human	immun

```

RESULT 2
Q7ZJN9 PRELIMINARY; PRT; 116 AA.
AC Q7ZJN9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08692A;
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol
RT integrase, and env gp41.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY214120; AA061840.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match 93.8%; Score 165; DB 15; Length 116;
Best Local Similarity 87.5%; Pred. No. 1.1e-17;
Matches 28; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLNSWCGKRLVCYTSVRWNET 32
Db 22 ALETLNQQLNSWCGKRLVCYTSVKWNET 53

RESULT 3
Q9IEB6 PRELIMINARY; PRT; 219 AA.
AC Q9IEB6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF14;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236404; CAB96252.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 219 219
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 92.6%; Score 163; DB 15; Length 219;
Best Local Similarity 87.5%; Pred. No. 4.5e-17;
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLNSWCGKRLVCYTSVRWNET 32
Db 47 ALETLNQQLNSWCGKRLVCYTSVKWNET 78

```

```

RESULT 4
Q9IHU7 PRELIMINARY; PRT; 124 AA.
AC Q9IHU7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM768;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF29237; AAF71914.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 15187 MW; 10SD515F114450F8 CRC64;

Query Match 91.5%; Score 161; DB 15; Length 124;
Best Local Similarity 84.4%; Pred. No. 5e-17;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLNSWCGKRLVCYTSVRWNET 32
Db 23 ALETLNQQLNSWCGKRLVCYTSVKWNET 54

RESULT 5
Q9IHV1 PRELIMINARY; PRT; 126 AA.
AC Q9IHV1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97ES203;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF29237; AAF71910.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 126 126

```

```
SQ SEQUENCE 126 AA; 15169 MW; 13FB101ECDF0DDDD CRC64;
Query Match 91.5%; Score 161; DB 15; Length 126;
Best Local Similarity 87.5%; Pred. No. 5.1e-17;
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 30 ALETLNQQLNSWGCGRGLVCYTSVKWNS 61

RESULT 6
Q9IHU9 PRELIMINARY; PRT; 130 AA.
AC Q9IHU9;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229235; AAF71912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1 130
FT NON_TER 130 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match 91.5%; Score 161; DB 15; Length 130;
Best Local Similarity 84.4%; Pred. No. 5.3e-17;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 30 ALETLNQQLNSWGCGRGLVCYTSVRWNET 61

RESULT 7
Q9DQL9 PRELIMINARY; PRT; 135 AA.
AC Q9DQL9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20584646; PubMed=11153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,
RA Vadiello J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).
```

```
DR EMBL; AF255939; AAC36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1 135
FT NON_TER 135 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match 91.5%; Score 161; DB 15; Length 135;
Best Local Similarity 84.4%; Pred. No. 5.5e-17;
Matches 27; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 38 ALETLNQQLNSWGCGRGLVCYTSVKWNST 69

RESULT 8
Q9IEB3 PRELIMINARY; PRT; 172 AA.
AC Q9IEB3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF99;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236407; CAB96255.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1 172
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20388 MW; 972C7AF2A7546B13 CRC64;

Query Match 91.5%; Score 161; DB 15; Length 172;
Best Local Similarity 84.4%; Pred. No. 7.1e-17;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLNSWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 ALETLNQQLNSWGCGRGLVCYTSVRWNET 57

RESULT 9
Q9IEC2 PRELIMINARY; PRT; 234 AA.
AC Q9IEC2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=BCF100;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236397; CAB96246.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 234 234
SQ SEQUENCE 234 AA; 27036 MW; 39B050B3F8555A8C CRC64;

Query Match 91.5%; Score 161; DB 15; Length 234;
Best Local Similarity 84.4%; Pred. No. 9.9e-17;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 48 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 79

RESULT 10
ID Q8Q7H0 PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=98CMA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282 (2002).
DR EMBL; AF383251; AAL98873.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 872 AA; 98344 MW; 91D37F60053827E3 CRC64;

Query Match 91.5%; Score 161; DB 15; Length 872;
Best Local Similarity 84.4%; Pred. No. 4e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 592 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 623

RESULT 11
Q8Q7F9
ID Q8Q7F9 PRELIMINARY; PRT; 882 AA.
AC Q8Q7F9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA100;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282 (2002).
DR EMBL; AF383252; AAL98884.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 627BCA5D0C09C6D4 CRC64;

Query Match 91.5%; Score 161; DB 15; Length 882;
Best Local Similarity 84.4%; Pred. No. 4.1e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 32
Db 603 ALETLLNQQLLSWGCGRGLVCYTSVRWNET 634

RESULT 12
Q8Q7G9
ID Q8Q7G9 PRELIMINARY; PRT; 887 AA.
AC Q8Q7G9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA8141;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282 (2002).
DR EMBL; AF383252; AAL98874.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 887 AA; 99559 MW; 88954F4ED76A314A CRC64;

Query Match 91.5%; Score 161; DB 15; Length 887;
Best Local Similarity 84.4%; Pred. No. 4.1e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```



```
QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 32
DB 608 ALETLLNQQLNSWGCGRGLVCYTSVKWNET 639

RESULT 13
Q9IE32 PRELIMINARY; PRT; 240 AA.
AC Q9IE32;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=BCFill;
RA Roques P., Robertson D., Sousquiere S., Damond F., Mauciere P.,
RA Depienne C., Brun-Vezinet F., Dornmont D., Simon F.O.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243365; CAB96335.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 240
SQ SEQUENCE 240 AA; 27723 MW; 1E206E8D491A4197 CRC64;

Query Match 90.9%; Score 160; DB 15; Length 240;
Best Local Similarity 84.4%; Pred. No. 1.9e-16;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 32
DB 50 ALETLLNQQLNSWGCGRGLVCYTSVKWNET 81

RESULT 14
Q9IHU8 PRELIMINARY; PRT; 125 AA.
AC Q9IHU8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=97CM756;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229236; AAF71913.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 125 AA; 15108 MW; EA2896C4B25FF342 CRC64;

Query Match 90.3%; Score 159; DB 15; Length 125;
Best Local Similarity 81.2%; Pred. No. 1e-16;
Matches 26; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 32
DB 27 ALETLLNQQLNSWGCGRGLVCYTSVKWNET 58

RESULT 15
Q9IEA5 PRELIMINARY; PRT; 216 AA.
AC Q9IEA5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=YBF26;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236415; CAB96263.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25003 MW; 0A5AC218BFA88932 CRC64;

Query Match 90.3%; Score 159; DB 15; Length 216;
Best Local Similarity 87.5%; Pred. No. 1.9e-16;
Matches 28; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 32
DB 32 ALETLLNQQLNSWGCGRGLVCYTSVRWNET 63

Search completed: May 7, 2004, 17:49:13
Job time : 40.605 secs
```

LINE MODE BLANK (USP70)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 60.9076 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362a-12

Perfect score: 176

Sequence: 1 ALETLNQQLNLINWGRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	100.0	32	2 AAW80470	Peptide d
2	170	96.6	32	2 AAW80469	Peptide d
3	169	96.0	32	2 AAW80471	Peptide d
4	165	93.8	200	3 AAY77373	HIV-1 gro
5	165	93.8	215	2 AAY09499	HIV-1 gro
6	165	93.8	215	2 AAY06983	Recombina
7	165	93.8	215	3 AAY77374	HIV-1 gro
8	165	93.8	245	2 AAY09493	HIV-1 gro
9	165	93.8	245	2 AAY06977	Recombina
10	165	93.8	245	3 AAY77369	HIV-1 gro
11	165	93.8	281	2 AAY09507	HIV-1 gro
12	165	93.8	373	2 AAY09495	HIV-1 gro
13	165	93.8	373	2 AAY06979	Recombina
14	165	93.8	460	2 AAY09500	HIV-1 gro
15	165	93.8	460	2 AAY06984	Recombina
16	165	93.8	460	3 AAY77375	HIV-1 gro
17	165	93.8	474	3 AAY77371	HIV-1 gro
18	165	93.8	488	2 AAY09504	HIV-1 gro
19	165	93.8	490	2 AAY09494	HIV-1 gro
20	165	93.8	490	2 AAY06978	Recombina
21	165	93.8	490	3 AAY77370	HIV-1 gro
22	165	93.8	526	2 AAY09505	HIV-1 gro
23	165	93.8	618	2 AAY09496	HIV-1 gro
24	165	93.8	618	2 AAY06980	Recombina
25	165	93.8	618	3 AAY77372	HIV-1 gro

ALIGNMENTS

RESULT 1  
AAW80470  
ID AAW80470 standard; peptide; 32 AA.

AC AAW80470;  
XX  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
PD 15-OCT-1998.  
XX  
PF 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX  
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX  
PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
XX immune deficiency virus of group O.  
PT  
XX Claim 6; Page 44; 55pp; French.

AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
(Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 32 AA;

Query Match 100.0%; Score 176; DB 2; Length 32;

Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32

RESULT 2  
AAW80469  
ID AAW80469 standard; peptide; 32 AA.  
AC AAW80469;  
XX  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
XX Peptide derived from a conserved sequence of group O human HIV.  
XX Group O human immune deficiency virus; HIV; detection; infection.  
XX  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
XX 15-OCT-1998.  
XX  
XX 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-2003 (revised)  
PR 25-MAR-2003 (revised)  
PR 28-JAN-1999 (first entry)  
XX  
XX  
XX Peptide derived from a conserved sequence of group O human HIV.  
XX Group O human immune deficiency virus; HIV; detection; infection.  
XX  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
XX 15-OCT-1998.  
XX  
XX 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
XX  
XX Claim 6; Page 44; 55pp; French.  
XX  
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
XX  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
XX  
XX Claim 6; Page 44; 55pp; French.  
XX  
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
XX  
XX  
SQ Sequence 32 AA;

Query Match 96.6%; Score 170; DB 2; Length 32;  
Best Local Similarity 96.9%; Pred. No. 1.5e-15;  
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32

RESULT 3  
AAW80471  
ID AAW80471 standard; peptide; 32 AA.  
AC AAW80471;  
XX  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX

XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
XX Group O human immune deficiency virus; HIV; detection; infection.  
XX  
XX Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
XX 15-OCT-1998.  
XX  
XX 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
XX  
XX Claim 6; Page 44; 55pp; French.  
XX  
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
XX  
XX  
SQ Sequence 32 AA;

Query Match 96.0%; Score 169; DB 2; Length 32;  
Best Local Similarity 93.8%; Pred. No. 2e-15;  
Matches 30; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 ALETLNQQLLNITWGCGRGLVCYTSVRWNET 32

RESULT 4  
AAW77373  
ID AAW77373 standard; protein; 200 AA.  
XX  
AC AAW77373;  
XX  
XX 22-MAY-2000 (first entry)  
XX  
XX  
DE HIV-1 group O env gp120/gp41 pGO-5 recombinant protein, SEQ ID NO:56.  
XX  
XX HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;  
XX immunoassay; positive control; affinity purification; therapeutic;  
XX Escherichia coli; antigen; synthetic gene construction; mutagen;  
XX deletion mutation.  
XX  
XX Human immunodeficiency virus 1; group O isolate HAM112.  
OS Synthetic.  
XX  
XX WO200004383-A2.  
XX  
XX 27-JAN-2000.  
XX  
XX 09-JUL-1999; 99WO-US015469.  
XX  
XX 14-JUL-1998; 98US-00115171.  
XX

(ABBO ) ABBOTT LAB.

Scheffel JW, Hackett JR, Tyner JD, Hickman RK;  
WPI; 2000-171290/15.  
N-PSDB; AAZ90284.

Novel monoclonal antibodies useful as positive control reagent for detecting human immunodeficiency virus infections and diagnosing, evaluating or prognosticating viral disease.

Example 3; Fig 9; 148pp; English.

The invention relates to anti-HIV-1 group O monoclonal antibodies, which may be used as positive control reagents in immunoassays to detect and differentiate HIV-1 infections. The invention also encompasses a monoclonal antibody which binds specifically to an HIV-1 group O antigen, which has no more than 15% cross reactivity to a corresponding antigen selected from HIV-1 group M antigens and HIV-2 antigens; and a method of using a monoclonal antibody as a positive control reagent in an immunoassay for the detection of anti HIV-1 group O antibodies. The monoclonal antibodies are useful as positive control reagents in immunoassays capable of detecting anti-HIV-1 group O antibodies. Such immunoassays involve coupling a monoclonal antibody with HIV group-1 antigen and detecting the antigen-antibody complex. The monoclonal antibodies of the invention would be used to ensure that the reagents provided to detect HIV-1 group O antibody were performing properly. The monoclonal antibodies may also can be immobilised on a matrix and used for affinity purification of specific HIV-1 group O-derived proteins from cell cultures or biological tissues. The monoclonal antibodies can also be used for generating chimeric antibodies for therapeutic use. Different synthetic, recombinant or purified antibodies which identify different epitopes of HIV antigens can be used in combination in assay to diagnose, evaluate, or prognosticate HIV disease condition. The monoclonal antibodies are also useful for differentiating HIV-1 Group O antigens from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent recombinant HIV-1 group O env antigens encoded by the synthetic genes AAZ90280-290286. The recombinant HIV-1 env proteins contain various deletions relative to the native HAM112 isolate env protein (AAV77376). The recombinant HIV-1 group O antigens were purified and used to screen hybridoma cultures

Query Match 93.8%; Score 165; DB 3; Length 200;  
Best Local Similarity 84.4%; Pred. No. 4.6e-14;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:  
Db 102 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 133

RESULT 5  
AAY09499  
ID AAY09499 standard; protein; 215 AA.  
AC AAY09499;  
XX AAY09499;  
XX AAY09499;  
DT 17-OCT-2003 (revised)  
DT 15-JUL-1999 (first entry)  
XX  
DE HIV-1 Group O env polypeptide pGO-8PL.  
XX  
XX HIV; human immunodeficiency virus; antigen; detection; antibody; differentiation; Group O; env; immunogen; immunoassay.  
KW  
XX Human immunodeficiency virus 1.  
OS  
XX WO9909179-A2.  
PN  
XX 25-FEB-1999.  
PD  
XX

17-AUG-1998; 98WO-US017014.  
XX  
PR 15-AUG-1997; 97US-00911824.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;  
XX WPI; 1999-190167/16.  
DR N-PSDB; AAX56078.  
XX  
PT New isolated HIV-1 Group O env polypeptides - used for the detection of anti-HIV antibodies and for the production of antibodies for use in detection, purification and therapy.  
XX  
PS Claim 17; Fig 5; 138pp; English.  
XX  
CC The present invention describes (A) an isolated HIV-1 Group O env polypeptide. Also described are: (1) an isolated HIV-1 Group O env polypeptide comprising an immunoreactive portion of a polypeptide as in (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1); (3) an antigen construct comprising a first HIV-1 Group O env polypeptide fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct comprising a fusion of at least one HIV-1 Group O env polypeptide with at least one HIV-1 Group M env polypeptide; (5) an antigen construct comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env polypeptide, and at least one additional HIV-1 polypeptide; (6) an antigen construct comprising a first HIV-2 env polypeptide fused to a second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a host cell transformed by an expression vector as in (8); and (10) an immunoassay kit for the detection of antibodies to HIV-1 comprising an antigen construct as in (3)-(6). The antigen constructs can be used for the detection of anti-HIV-1 antibodies in test samples. They can also be used as immunogens to produce antibodies. The antibodies can be used to purify HIV polypeptides, for therapy and for detection of HIV polypeptides. (Updated on 17-OCT-2003 to standardise OS field)

Query Match 93.8%; Score 165; DB 2; Length 215;  
Best Local Similarity 84.4%; Pred. No. 5e-14;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:  
Db 117 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 148

RESULT 6  
AAY06983  
ID AAY06983 standard; protein; 215 AA.  
XX AAY06983;  
XX AAY06983;  
DT 06-JUL-1999 (first entry)  
XX  
DE Recombinant pGO-8PL protein.  
XX  
KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening; antibody; assay.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
FH Location/Qualifiers  
FT Protein 2..46  
FT Peptide /note= "gp120 sequence"  
FT Peptide 47..245  
FT Peptide /note= "gp41 sequence"  
XX  
XX WO9909410-A2.  
XX

(ABBO ) ABBOTT LAB.

Scheffel JW, Hackett JR, Tyner JD, Hickman RK;  
WPI; 2000-171290/15.  
N-PSDB; AAZ90284.

Novel monoclonal antibodies useful as positive control reagent for detecting human immunodeficiency virus infections and diagnosing, evaluating or prognosticating viral disease.

Example 3; Fig 9; 148pp; English.

The invention relates to anti-HIV-1 group O monoclonal antibodies, which may be used as positive control reagents in immunoassays to detect and differentiate HIV-1 infections. The invention also encompasses a monoclonal antibody which binds specifically to an HIV-1 group O antigen, which has no more than 15% cross reactivity to a corresponding antigen selected from HIV-1 group M antigens and HIV-2 antigens; and a method of using a monoclonal antibody as a positive control reagent in an immunoassay for the detection of anti HIV-1 group O antibodies. The monoclonal antibodies are useful as positive control reagents in immunoassays capable of detecting anti-HIV-1 group O antibodies. Such immunoassays involve coupling a monoclonal antibody with HIV group-1 antigen and detecting the antigen-antibody complex. The monoclonal antibodies of the invention would be used to ensure that the reagents provided to detect HIV-1 group O antibody were performing properly. The monoclonal antibodies may also can be immobilised on a matrix and used for affinity purification of specific HIV-1 group O-derived proteins from cell cultures or biological tissues. The monoclonal antibodies can also be used for generating chimeric antibodies for therapeutic use. Different synthetic, recombinant or purified antibodies which identify different epitopes of HIV antigens can be used in combination in assay to diagnose, evaluate, or prognosticate HIV disease condition. The monoclonal antibodies are also useful for differentiating HIV-1 Group O antigens from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent recombinant HIV-1 group O env antigens encoded by the synthetic genes AAZ90280-290286. The recombinant HIV-1 env proteins contain various deletions relative to the native HAM112 isolate env protein (AAV77376). The recombinant HIV-1 group O antigens were purified and used to screen hybridoma cultures

Query Match 93.8%; Score 165; DB 3; Length 200;  
Best Local Similarity 84.4%; Pred. No. 4.6e-14;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:  
Db 102 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 133

RESULT 5  
AAY09499  
ID AAY09499 standard; protein; 215 AA.  
AC AAY09499;  
XX AAY09499;  
XX AAY09499;  
DT 17-OCT-2003 (revised)  
DT 15-JUL-1999 (first entry)  
XX  
DE HIV-1 Group O env polypeptide pGO-8PL.  
XX  
XX HIV; human immunodeficiency virus; antigen; detection; antibody; differentiation; Group O; env; immunogen; immunoassay.  
KW  
XX Human immunodeficiency virus 1.  
OS  
XX WO9909179-A2.  
PN  
XX 25-FEB-1999.  
PD  
XX

17-AUG-1998; 98WO-US017014.  
XX  
PR 15-AUG-1997; 97US-00911824.  
XX  
XX (ABBO ) ABBOTT LAB.  
XX  
PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;  
XX WPI; 1999-190167/16.  
DR N-PSDB; AAX56078.  
XX  
PT New isolated HIV-1 Group O env polypeptides - used for the detection of anti-HIV antibodies and for the production of antibodies for use in detection, purification and therapy.  
XX  
PS Claim 17; Fig 5; 138pp; English.  
XX  
CC The present invention describes (A) an isolated HIV-1 Group O env polypeptide. Also described are: (1) an isolated HIV-1 Group O env polypeptide comprising an immunoreactive portion of a polypeptide as in (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1); (3) an antigen construct comprising a first HIV-1 Group O env polypeptide fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct comprising a fusion of at least one HIV-1 Group O env polypeptide with at least one HIV-1 Group M env polypeptide; (5) an antigen construct comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env polypeptide, and at least one additional HIV-1 polypeptide; (6) an antigen construct comprising a first HIV-2 env polypeptide fused to a second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a host cell transformed by an expression vector as in (8); and (10) an immunoassay kit for the detection of antibodies to HIV-1 comprising an antigen construct as in (3)-(6). The antigen constructs can be used for the detection of anti-HIV-1 antibodies in test samples. They can also be used as immunogens to produce antibodies. The antibodies can be used to purify HIV polypeptides, for therapy and for detection of HIV polypeptides. (Updated on 17-OCT-2003 to standardise OS field)

Query Match 93.8%; Score 165; DB 2; Length 215;  
Best Local Similarity 84.4%; Pred. No. 5e-14;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:  
Db 117 ALETLQNQLNLINWCGRGLVCYTTSVRWNET 148

RESULT 6  
AAY06983  
ID AAY06983 standard; protein; 215 AA.  
XX AAY06983;  
XX AAY06983;  
DT 06-JUL-1999 (first entry)  
XX  
DE Recombinant pGO-8PL protein.  
XX  
KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening; antibody; assay.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
FH Location/Qualifiers  
FT Protein 2..46  
FT Peptide /note= "gp120 sequence"  
FT Peptide 47..245  
FT Peptide /note= "gp41 sequence"  
XX  
XX WO9909410-A2.  
XX









```

PD 25-FEB-1999.
XX
XX
PF 17-AUG-1998; 98WO-US017014.
XX
XX
PR 15-AUG-1997; 97US-00911824.
XX
XX
PA (ABBO ) ABBOTT LAB.
XX
XX
PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
XX
XX
DR WPI; 1999-190167/16.
XX
XX
DR N-PSDB; AAX56076.
XX
XX
PT New isolated HIV-1 Group O env polypeptides - used for the detection of
XX
XX
PT anti-HIV antibodies and for the production of antibodies for use in
XX
XX
PT detection, purification and therapy.
XX
XX
PS Claim 15; Fig 9; 138pp; English.
XX
XX
CC The present invention describes (A) an isolated HIV-1 Group O env
XX
XX
CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env
XX
XX
CC polypeptide comprising an immunoreactive portion of a polypeptide as in
XX
XX
CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
XX
XX
CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
XX
XX
CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
XX
XX
CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at
XX
XX
CC least one HIV-1 Group M env polypeptide; (5) an antigen construct
XX
XX
CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
XX
XX
CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an
XX
XX
CC antigen construct comprising a first HIV-2 env polypeptide fused to a
XX
XX
CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
XX
XX
CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
XX
XX
CC host cell transformed by an expression vector as in (8); and (10) an
XX
XX
CC immunoassay kit for the detection of antibodies to HIV-1 comprising an
XX
XX
CC antigen construct as in (3)-(6). The antigen constructs can be used for
XX
XX
CC the detection of anti-HIV-1 antibodies in test samples. They can also be
XX
XX
CC used as immunogens to produce antibodies. The antibodies can be used to
XX
XX
CC purify HIV polypeptides, for therapy and for detection of HIV
XX
XX
CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX
SQ Sequence 373 AA;
Query Match 93.8%; Score 165; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 8.9e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLNWGCGRILVCYTSVRWNET 32
Db 117 ALETLLNQQLLNWGCGRILVCYTSVRWNET 148
RESULT 13
RAY06979
ID AAY06979 standard; protein; 373 AA.
XX
XX
AC AAY06979;
XX
XX
DT 06-JUL-1999 (first entry)
XX
XX
DE Recombinant pGO-11PL protein.
XX
XX
KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;
XX
XX
KW antibody; assay.
XX
XX
OS Synthetic.
XX
XX
OS Human immunodeficiency virus 1.
XX
XX
FH Key Location/Qualifiers
XX
XX
FT Protein 2..46
XX
XX
FT Peptide /note= "gpi20 sequence"
XX
XX
FT /note= "gp41 sequence"
XX
XX
XX

```

---

```

PN WO9909410-A2.
XX
XX
PD 25-FEB-1999.
XX
XX
PF 07-AUG-1998; 98WO-US016506.
XX
XX
PR 15-AUG-1997; 97US-00912129.
XX
XX
PA (ABBO ) ABBOTT LAB.
XX
XX
PI Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
XX
XX
PI Golden AM, Brennan CA, Devare SG;
XX
XX
DR WPI; 1999-190224/16.
XX
XX
DR N-PSDB; AAX37191.
XX
XX
PT New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
XX
XX
PT be used in field assay, requiring no electricity and less specialised
XX
XX
PT equipment.
XX
XX
PS Claim 1; Fig 9; 104pp; English.
XX
XX
CC The invention relates to a rapid assay for simultaneous detection and
XX
XX
CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
XX
XX
CC method comprises (a) contacting the sample with a strip containing at
XX
XX
CC least one immobilised capture reagent per analyte and on which the sample
XX
XX
CC moves from the proximal to the distal end by capillary action, under
XX
XX
CC conditions sufficient to form capture reagent/analyte complexes, and (b)
XX
XX
CC determining the presence of analyte(s) by detecting a visible colour
XX
XX
CC change at the capture reagent site on the strip wherein the capture
XX
XX
CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
XX
XX
CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
XX
XX
CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in
XX
XX
CC AAY06981. The invention is used to screen patients for antibodies to HIV-
XX
XX
CC 1 types O and M, and HIV-2. The invention will be particularly useful in
XX
XX
CC places and situation where equipment and/or electricity is not available.
XX
XX
CC The invention provides a screening method which is faster and requires
XX
XX
CC less equipment than prior art methods. The present sequence represents a
XX
XX
CC amino acid sequence of the recombinant pGO-11PL recombinant protein which
XX
XX
CC acts as a capture reagent for HIV-1 group O
XX
XX
SQ Sequence 373 AA;
Query Match 93.8%; Score 165; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 8.9e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLNWGCGRILVCYTSVRWNET 32
Db 117 ALETLLNQQLLNWGCGRILVCYTSVRWNET 148
RESULT 14
RAY09500
ID AAY09500 standard; protein; 460 AA.
XX
XX
AC AAY09500;
XX
XX
DT 17-OCT-2003 (revised)
XX
XX
DT 15-JUL-1999 (first entry)
XX
XX
DE HIV-1 Group O env polypeptide pGO-8CKS.
XX
XX
KW HIV; human immunodeficiency virus; antigen; detection; antibody;
XX
XX
KW differentiation; Group O; env; immunogen; immunoassay.
XX
XX
OS Human immunodeficiency virus 1.
XX
XX
OS
XX
XX
FH Key Location/Qualifiers
XX
XX
FT Protein 2..46
XX
XX
FT Peptide /note= "gpi20 sequence"
XX
XX
FT /note= "gp41 sequence"
XX
XX
XX

```

```
XX 15-AUG-1997; 97US-00911824.
XX (ABBO ) ABBOTT LAB.
XX Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
XX WPI; 1999-190167/16.
XX N-PSDB; AAX56079.
XX New isolated HIV-1 Group O env polypeptides - used for the detection of
XX anti-HIV antibodies and for the production of antibodies for use in
XX detection, purification and therapy.
XX Example 3; Fig 6; 138pp; English.
XX The present invention describes (A) an isolated HIV-1 Group O env
XX polypeptide. Also described are: (1) an isolated HIV-1 Group O env
XX polypeptide comprising an immunoreactive portion of a polypeptide as in
XX (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
XX (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
XX fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
XX comprising a fusion of at least one HIV-1 Group O env polypeptide with at
XX least one HIV-1 Group M env polypeptide; (5) an antigen construct
XX comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
XX polypeptide, and at least one additional HIV-1 polypeptide; (6) an
XX antigen construct comprising a first HIV-2 env polypeptide fused to a
XX second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
XX in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
XX host cell transformed by an expression vector as in (8); and (10) an
XX immunoassay kit for the detection of antibodies to HIV-1 comprising an
XX antigen construct as in (3)-(6). The antigen constructs can be used for
XX the detection of anti-HIV-1 antibodies in test samples. They can also be
XX used as immunogens to produce antibodies. The antibodies can be used to
XX purify HIV polypeptides, for therapy and for detection of HIV
XX polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX Sequence 460 AA;
XX
XX Query Match 93.8%; Score 165; DB 2; Length 460;
XX Best Local Similarity 84.4%; Pred. No. 1.1e-13;
XX Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLNQQLLNWCGKRLVCYTSVRWNET 32
Db 362 ALETLNQQLLNWCGKRLVCYTSVKWNET 393
RESULT 15
AAY06984
XX AAY06984 standard; protein; 460 AA.
XX AAY06984;
XX 06-JUL-1999 (first entry)
XX Recombinant pGO-8CKS protein.
XX HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;
XX antibody; assay.
XX Synthetic.
XX Human immunodeficiency virus 1.
XX Key Location/Qualifiers
XX Protein 1..246
XX /note= "CKS sequence"
XX Protein 247..291
XX /note= "gp120 sequence"
XX Protein 292..460
XX /note= "gp41 sequence"
XX WO9909410-A2.
XX
XX 25-FEB-1999.
XX 07-AUG-1998; 98WO-US016506.
XX 15-AUG-1997; 97US-00912129.
XX (ABBO ) ABBOTT LAB.
XX Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
XX Golden AM, Brennan CA, Devare SG;
XX WPI; 1999-190224/16.
XX N-PSDB; AAX37194.
XX New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
XX be used in field assay, requiring no electricity and less specialised
XX equipment.
XX Claim 1; Fig 6; 104pp; English.
XX The invention relates to a rapid assay for simultaneous detection and
XX differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
XX method comprises (a) contacting the sample with a strip containing at
XX least one immobilised capture reagent per analyte and on which the sample
XX moves from the proximal to the distal end by capillary action, under
XX conditions sufficient to form capture reagent/analyte complexes, and (b)
XX determining the presence of analyte(s) by detecting a visible colour
XX change at the capture reagent site on the strip wherein the capture
XX reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
XX and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
XX in AAY06982; and that for HIV-2 comprises the polypeptide shown in
XX AAY06981. The invention is used to screen patients for antibodies to HIV-
XX 1 types O and M, and HIV-2. The invention will be particularly useful in
XX places and situation where equipment and/or electricity is not available.
XX The invention provides a screening method which is faster and requires
XX less equipment than prior art methods. The present sequence represents a
XX amino acid sequence of the recombinant pGO-8CKS protein which acts as a
XX capture reagent for HIV-1 group O
XX
XX Sequence 460 AA;
XX
XX Query Match 93.8%; Score 165; DB 2; Length 460;
XX Best Local Similarity 84.4%; Pred. No. 1.1e-13;
XX Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLNQQLLNWCGKRLVCYTSVRWNET 32
Db 362 ALETLNQQLLNWCGKRLVCYTSVKWNET 393
Search completed: May 7, 2004, 17:42:44
Job time : 60.9076 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 17.6134 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-12

Perfect score: 176

Sequence: 1 ALETLNQQLLNWGRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pdp.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pdp.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pdp.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pdp.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTJUS\_COMB.pdp.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	93.8	215	2	US-08-912-129A-58
2	165	93.8	245	2	US-08-912-129A-48
3	165	93.8	373	2	US-08-912-129A-52
4	165	93.8	460	2	US-08-912-129A-60
5	165	93.8	490	2	US-08-912-129A-50
6	165	93.8	618	2	US-08-912-129A-54
7	165	93.8	873	2	US-08-912-129A-61
8	162	92.0	40	3	US-08-894-699-39
9	162	92.0	40	3	US-09-444-410-39
10	162	92.0	116	4	US-09-462-917A-20
11	162	92.0	356	1	US-08-602-713-12
12	162	92.0	356	3	US-08-989-493-12
13	162	92.0	356	4	US-09-610-271-12
14	161	91.5	113	4	US-09-462-917A-2
15	161	91.5	117	4	US-09-462-917A-6
16	160	90.9	41	3	US-08-894-699-67
17	160	90.9	41	3	US-09-444-410-67
18	160	90.9	715	4	US-08-462-917A-134
19	159	90.3	37	4	US-08-817-441-94
20	159	90.3	104	4	US-08-817-441-100
21	159	90.3	113	4	US-09-462-917A-10
22	159	90.3	113	4	US-09-462-917A-12
23	158	89.8	40	3	US-08-894-699-37
24	158	89.8	40	3	US-08-894-699-42
25	158	89.8	40	3	US-09-444-410-37
26	158	89.8	40	3	US-09-444-410-42
27	158	89.8	110	4	US-09-462-917A-30

28	157	89.2	33	3	US-09-433-428D-10	Sequence 10, Appl
29	157	89.2	40	3	US-08-894-699-40	Sequence 40, Appl
30	157	89.2	40	3	US-09-444-410-40	Sequence 40, Appl
31	157	89.2	113	4	US-09-462-917A-28	Sequence 28, Appl
32	156	88.6	33	3	US-09-433-428D-4	Sequence 4, Appl
33	156	88.6	33	3	US-09-433-428D-15	Sequence 15, Appl
34	156	88.6	42	3	US-08-894-699-66	Sequence 66, Appl
35	156	88.6	42	3	US-09-444-410-66	Sequence 66, Appl
36	156	88.6	113	4	US-09-462-917A-18	Sequence 18, Appl
37	156	88.6	116	4	US-09-462-917A-22	Sequence 22, Appl
38	155	88.1	33	3	US-09-433-428D-5	Sequence 5, Appl
39	155	88.1	33	3	US-09-433-428D-12	Sequence 12, Appl
40	155	88.1	37	4	US-08-817-441-86	Sequence 86, Appl
41	155	88.1	200	3	US-08-965-056-104	Sequence 104, App
42	155	88.1	862	4	US-09-206-551-15	Sequence 15, Appl
43	154	87.5	33	3	US-09-433-428D-8	Sequence 8, Appl
44	154	87.5	115	4	US-09-462-917A-24	Sequence 24, Appl
45	153	86.9	33	3	US-09-433-428D-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1  
US-08-912-129A-58  
; Sequence 58, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.  
; APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.  
; APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.  
; APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.  
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS (Windows 95)  
; SOFTWARE: Microsoft Word (ASCII format output)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,129A  
; FILING DATE: 15-AUG-1997  
; CLASSIFICATION: 436  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dancigers, Andreas M.  
; REGISTRATION NUMBER: 32,652  
; REFERENCE/DOCKET NUMBER: 6109.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847-937-9803  
; TELEFAX: 847-938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 215 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-912-129A-58

```

; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancakers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-52

Query Match 93.8%; Score 165; DB 2; Length 373;
Best Local Similarity 84.4%; Pred.No. 1.2e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQOQLNLINWGCRGLVCYTSVRNET 32
   |||||:|||||:|||||:|||||:|||||
Db 117 ALETLQOQLNLINWGCKRGLICYTSVKWNET 148

RESULT 4
US-08-912-129A-60
; Sequence 60, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories

```

```
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS (Windows 95)
SOFTWARE: Microsoft Word (ASCII format output)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,129A
FILING DATE: 15-AUG-1997
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M.
REGISTRATION NUMBER: 32,652
REFERENCE/DOCKET NUMBER: 6109.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-937-9803
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-912-129A-60

Query Match 93.8%; Score 165; DB 2; Length 460;
Best Local Similarity 84.4%; Pred. No. 1.5e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNWCGRLVCYTSVRWNET 32
Db 362 ALETLLNQQLLNWCGRLVCYTSVRWNET 393

RESULT 5
US-08-912-129A-50
Sequence 50, Application US/08912129A
Patent No. 5922533
GENERAL INFORMATION:
APPLICANT: VALLARI, ANADRUZELA S.
APPLICANT: HACKETT, JOHN JR.
APPLICANT: HICKMAN, ROBERT K.
APPLICANT: VARITEK, VINCENT A. JR.
APPLICANT: NECKLAWS, ELIZABETH A.
APPLICANT: GOLDEN, ALAN M.
APPLICANT: BRENNAN, CATHERINE A.
APPLICANT: DEVARE, SUSHIL G.
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS (Windows 95)
SOFTWARE: Microsoft Word (ASCII format output)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,129A
FILING DATE: 15-AUG-1997
```

```
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M.
REGISTRATION NUMBER: 32,652
REFERENCE/DOCKET NUMBER: 6109.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-937-9803
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-912-129A-50

Query Match 93.8%; Score 165; DB 2; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.6e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNWCGRLVCYTSVRWNET 32
Db 362 ALETLLNQQLLNWCGRLVCYTSVRWNET 393

RESULT 6
US-08-912-129A-54
Sequence 54, Application US/08912129A
Patent No. 5922533
GENERAL INFORMATION:
APPLICANT: VALLARI, ANADRUZELA S.
APPLICANT: HACKETT, JOHN JR.
APPLICANT: HICKMAN, ROBERT K.
APPLICANT: VARITEK, VINCENT A. JR.
APPLICANT: NECKLAWS, ELIZABETH A.
APPLICANT: GOLDEN, ALAN M.
APPLICANT: BRENNAN, CATHERINE A.
APPLICANT: DEVARE, SUSHIL G.
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS (Windows 95)
SOFTWARE: Microsoft Word (ASCII format output)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,129A
FILING DATE: 15-AUG-1997
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M.
REGISTRATION NUMBER: 32,652
REFERENCE/DOCKET NUMBER: 6109.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-937-9803
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 54:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-129A-54

Query Match          93.8%; Score 165; DB 2; Length 618;
Best Local Similarity 84.4%; Pred. No. 2e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ALETLNQQLLNINWCGKRLVCYTSVRWNET 32
Db      362 ALETLNQQLLNINWCGKRLVCYTSVRWNET 393

RESULT 7
US-08-912-129A-61
; Sequence 61, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912.129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancikers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109 US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 873 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-912-129A-61

Query Match          93.8%; Score 165; DB 2; Length 873;
Best Local Similarity 84.4%; Pred. No. 3e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ALETLNQQLLNINWCGKRLVCYTSVRWNET 32
```

```
Db      591 ALETLNQQLLNINWCGKRLVCYTSVRWNET 622

RESULT 8
US-08-894-699-39
; Sequence 39, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-OPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-699-39

Query Match          92.0%; Score 162; DB 3; Length 40;
Best Local Similarity 87.5%; Pred. No. 2.6e-17;
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ALETLNQQLLNINWCGKRLVCYTSVRWNET 32
Db      9 ALETLNQQLLNINWCGKRLVCYTSVRWNET 40

RESULT 9
US-09-444-410-39
; Sequence 39, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
```

APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 92.0%; Score 162; DB 3; Length 40;  
Best Local Similarity 87.5%; Pred. No. 2.6e-17;  
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 9 ALETLNQQLLNWGCGRGLVCYTSVRWNET 40

RESULT 10  
US-09-462-917A-20  
Sequence 20, Application US/09462917A  
Patent No. 6511801  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
APPLICANT: Vanden Haesevelde, Marlen  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014 11362.0014.NPUS00  
CURRENT APPLICATION NUMBER: US/09/462,917A  
CURRENT FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 20

LENGTH: 116  
TYPE: PRT  
ORGANISM: Human  
US-09-462-917A-20

Query Match 92.0%; Score 162; DB 4; Length 116;  
Best Local Similarity 84.4%; Pred. No. 8.7e-17;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 29 ALETLNQQLLNWGCGRGLVCYTSVRWNET 60

RESULT 11  
US-08-602-713-12  
Sequence 12, Application US/08602713  
Patent No. 5798205  
GENERAL INFORMATION:  
APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;  
APPLICANT: G itler, Lutz G.; Eberle, Josef; Kaptue, Lazare;  
APPLICANT: Zekeng, L opold Achengu  
TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use  
TITLE OF INVENTION: (MVP-2901/94)  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,713  
FILING DATE: 16-FEBRUARY-1996  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 195 05 262  
FILING DATE: 16-FEBRUARY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5798205man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LEDE 203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
US-08-602-713-12

Query Match 92.0%; Score 162; DB 1; Length 356;  
Best Local Similarity 81.2%; Pred. No. 3.1e-16;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWGCGRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||  
Db 312 ALETLNQQLLNWGCGRGLVCYTSVRWNET 343

RESULT 12  
US-08-989-493-12  
Sequence 12, Application US/08989493

```

;
; COMPUTER READABLE FORM:
; MEDIAN TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,271
; FILING DATE: 06-Jul-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,713
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6548635man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LEDER 203
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 698-9200
; TELEFAX: (212) 838-3884
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
;
; MOLECULE TYPE: Protein
; FRAGMENT TYPE: internal
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
;
US-09-610-271-12
;
Query Match 92.0%; Score 162; DB 4; Length 356;
Best Local Similarity 81.2%; Pred. No. 3.1e-16;
Matches 26; Conservative 6; Mismatches 0; Indels 0;

Qy 1 ALETLNQQLLNWGCGRIVCYTSVRWNET 32
| | | | | | | | | | | | | | | | | | | | | |
Db 312 ALETLNQQLLNWGCGRIVCYTSVRWNET 343

RESULT 14
US-09-462-917A-2
; Sequence 2, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NPUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
US-09-462-917A-2
;
Query Match 91.5%; Score 161; DB 4; Length 113;
Best Local Similarity 81.2%; Pred. No. 1.2e-16;
Matches 26; Conservative 6; Mismatches 0; Indels 0;

Qy 1 ALETLNQQLLNWGCGRIVCYTSVRWNET 32
| | | | | | | | | | | | | | | | | | | | | |
Db 29 ALETLNQQLLNWGCGRIVCYTSVRWNET 60

RESULT 15
US-09-462-917A-6

```



```
; Sequence 6, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NPUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Human
US-09-462-917A-6
```

```
Query Match          91.5%; Score 161; DB 4; Length 117;
Best Local Similarity 84.4%; Pred. No. 1.2e-16;
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ALETLLNQQLLNWGCGRGLVCYTSVRWNET 32
        |||||:|||||:|||||:|||||:|||||
Db      29 ALETLLNQQLLNWGCGRGLVCYTSVKWNT 60
```

Search completed: May 7, 2004, 17:53:26  
Job time : 17.6134 secs

THE FINE DUTY (US10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 44.7731 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-12

Perfect score: 176

Sequence: 1 ALETLLNQQLLNWCGRLVCYTSVRMNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pap.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pap.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.pap.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb.pap.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.pap.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb.pap.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.pap.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb.pap.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pdb.pap.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pdb.pap.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pdb.pap.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.pap.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pdb.pap.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pdb.pap.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pdb.pap.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.pap.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.pap.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	93.8	215	8	US-08-911-824-58
2	165	93.8	245	8	US-08-911-824-48
3	165	93.8	281	8	US-08-911-824-120
4	165	93.8	373	8	US-08-911-824-52
5	165	93.8	460	8	US-08-911-824-60
6	165	93.8	488	8	US-08-911-824-95
7	165	93.8	490	8	US-08-911-824-50
8	165	93.8	526	8	US-08-911-824-97
9	165	93.8	618	8	US-08-911-824-54
10	165	93.8	706	8	US-08-911-824-93
11	165	93.8	736	8	US-08-911-824-91
12	165	93.8	873	8	US-08-911-824-61
13	162	92.0	116	14	US-10-320-786-20
14	162	92.0	356	14	US-10-357-400-12
15	161	91.5	113	14	US-10-320-786-2

16	161	91.5	117	14	US-10-320-786-6	Sequence 6, Appli
17	160	90.9	715	14	US-10-320-786-134	Sequence 134, App
18	159	90.3	37	14	US-10-026-741-94	Sequence 94, Appl
19	159	90.3	104	14	US-10-026-741-100	Sequence 100, App
20	159	90.3	113	14	US-10-320-786-10	Sequence 10, Appl
21	159	90.3	113	14	US-10-320-786-12	Sequence 12, Appl
22	158	89.8	110	14	US-10-320-786-30	Sequence 30, Appl
23	157	89.2	113	14	US-10-320-786-28	Sequence 28, Appl
24	156	88.6	113	14	US-10-320-786-18	Sequence 18, Appl
25	156	88.6	116	14	US-10-320-786-22	Sequence 22, Appl
26	155	88.1	37	14	US-10-026-741-86	Sequence 86, Appl
27	155	88.1	200	9	US-09-854-816-104	Sequence 104, App
28	155	88.1	862	14	US-10-369-294-15	Sequence 15, Appl
29	154	87.5	115	14	US-10-320-786-24	Sequence 24, Appl
30	152	86.4	113	14	US-10-320-786-40	Sequence 40, Appl
31	151	85.8	35	14	US-10-026-741-101	Sequence 101, App
32	151	85.8	37	14	US-10-026-741-7	Sequence 7, Appli
33	151	85.8	37	14	US-10-026-741-88	Sequence 88, Appl
34	151	85.8	37	14	US-10-026-741-90	Sequence 90, Appl
35	151	85.8	110	14	US-10-320-786-36	Sequence 36, Appl
36	151	85.8	351	14	US-10-026-741-47	Sequence 47, Appl
37	151	85.8	877	14	US-10-026-741-102	Sequence 102, App
38	150	85.2	110	14	US-10-320-786-14	Sequence 14, Appl
39	149	84.7	110	14	US-10-320-786-16	Sequence 16, Appl
40	146	83.0	35	9	US-09-886-156-62	Sequence 62, Appl
41	146	83.0	35	9	US-09-886-150-62	Sequence 62, Appl
42	146	83.0	35	10	US-09-886-149-62	Sequence 62, Appl
43	146	83.0	35	10	US-09-886-159-62	Sequence 62, Appl
44	146	83.0	35	14	US-10-326-090-62	Sequence 62, Appl
45	146	83.0	146	13	US-10-000-321-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 93.8% Score 165; DB 8; Length 215;

Best Local Similarity 84.4%; Pred. No. 5.2e-15; Indels 0; Gaps 0;  
Matches 27; Conservative 5; Mismatches 0;

Qy 1 ALETLLNQQLLNWCGRLVCYTSVRMNET 32  
|||||:|||||:|||||:|||||:|||||:  
Db 117 ALETLLNQQLLNWCGRLVCYTSVRMNET 148

RESULT 2

US-08-911-824-48



```

; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-14PL
US-08-911-824-95

Query Match          93.8%; Score 165; DB 8; Length 488;
Best Local Similarity 84.4%; Pred. No. 1.2e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 117 ALETLNQQLLNINWGCRGLVCYTSVRWNET 148

RESULT 7
US-08-911-824-50
; Sequence 50, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-50

Query Match          93.8%; Score 165; DB 8; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.2e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393

RESULT 8
US-08-911-824-97
; Sequence 97, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS
US-08-911-824-97

Query Match          93.8%; Score 165; DB 8; Length 526;
Best Local Similarity 84.4%; Pred. No. 1.3e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393

RESULT 9
US-08-911-824-54
; Sequence 54, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

Query Match          93.8%; Score 165; DB 8; Length 618;
Best Local Similarity 84.4%; Pred. No. 1.5e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393

RESULT 10
US-08-911-824-93
; Sequence 93, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-93

Query Match          93.8%; Score 165; DB 8; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.2e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393
```

```

; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS
US-08-911-824-97

Query Match          93.8%; Score 165; DB 8; Length 526;
Best Local Similarity 84.4%; Pred. No. 1.3e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393

RESULT 9
US-08-911-824-54
; Sequence 54, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

Query Match          93.8%; Score 165; DB 8; Length 618;
Best Local Similarity 84.4%; Pred. No. 1.5e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393

RESULT 10
US-08-911-824-93
; Sequence 93, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-93

Query Match          93.8%; Score 165; DB 8; Length 490;
Best Local Similarity 84.4%; Pred. No. 1.2e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWGCRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||
DB 362 ALETLNQQLLNINWGCRGLVCYTSVRWNET 393
```

```
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165-US-01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-13CKS
US-08-911-824-93

Query Match          93.8%; Score 165; DB 8; Length 706;
Best Local Similarity 84.4%; Pred. No. 1.8e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 608 ALETLNQQLLNWCGRGLVCYTSVRWNET 639

RESULT 11
US-08-911-824-91
; Sequence 91, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165-US-01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-12CKS
US-08-911-824-91

Query Match          93.8%; Score 165; DB 8; Length 736;
Best Local Similarity 84.4%; Pred. No. 1.8e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 608 ALETLNQQLLNWCGRGLVCYTSVRWNET 639

RESULT 12
US-08-911-824-61
; Sequence 61, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
```

```
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165-US-01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: HIV-1 Group O isolate HAM112
US-08-911-824-61

Query Match          93.8%; Score 165; DB 8; Length 873;
Best Local Similarity 84.4%; Pred. No. 2.2e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 591 ALETLNQQLLNWCGRGLVCYTSVRWNET 622

RESULT 13
US-10-320-786-20
; Sequence 20, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014-1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; OTHER INFORMATION:
US-10-320-786-20

Query Match          92.0%; Score 162; DB 14; Length 116;
Best Local Similarity 84.4%; Pred. No. 7.3e-15;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNWCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 29 ALETLNQQLLNWCGRGLVCYTSVRWNET 60

RESULT 14
US-10-357-400-12
; Sequence 12, Application US/10357400
; Publication No. US20030147917A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
; Grottel, Lutz G.; Eberle, Josef; Kapue, Lazare;
; Zekeng, L opold Achengui
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use
; (MVP-2901/94)
```

```

RESULT 15
US-10-320-786-2
; Sequence 2, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-2

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 13.4454 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-12

Perfect score: 176

Sequence: 1 ALETLLNQQLLNIGCRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283365

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	90.3	104	2	S52930
2	155	88.1	863	2	A53034
3	151	85.8	877	2	S49137
4	114	64.8	854	1	VCLJSH
5	111	63.1	357	2	S21990
6	111	63.1	358	2	S22002
7	111	63.1	454	2	B41621
8	110	62.5	358	2	S22000
9	110	62.5	358	2	S70417
10	109	61.9	443	2	C41621
11	109	61.9	853	2	S54384
12	109	61.9	855	1	VCLJZR
13	108	61.4	357	2	S22006
14	108	61.4	357	2	S21994
15	108	61.4	357	2	S22004
16	108	61.4	357	2	S21996
17	108	61.4	357	2	S21992
18	108	61.4	358	2	S21998
19	108	61.4	445	2	A41621
20	108	61.4	843	1	H44001
21	108	61.4	852	1	VCLJBR
22	108	61.4	852	2	T12016
23	108	61.4	854	2	S13288
24	108	61.4	854	1	VCLJAZ
25	108	61.4	856	1	VCLJH3
26	108	61.4	856	1	VCLJVL
27	108	61.4	856	1	VCLJ3W
28	108	61.4	861	1	VCLJLV
29	108	61.4	861	1	VCLJSC

30	106	60.2	856	1	A44963
31	105	59.7	846	1	VCLJND
32	105	59.7	847	2	T09448
33	105	59.7	847	2	S13289
34	104	59.1	859	1	VCLJMN
35	104	59.1	868	1	VCLJHA
36	102	58.0	729	1	VCLJXX
37	102	58.0	861	1	VCLJKB
38	101	57.4	851	2	S33985
39	96	54.5	859	2	T01672
40	95.5	54.3	855	2	A45713
41	93	52.8	732	2	S46352
42	89.5	50.9	859	1	VCLJST
43	89.5	50.9	859	2	S24571
44	89.5	50.9	885	2	S04322
45	89.5	50.9	886	2	T11555

ALIGNMENTS

RESULT 1

S52930

GP41 ENV protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C:Accession: S52930

R:Cohen, J.H.M.; Guetard, D.; Philbert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de

submitted to the EMBL Data Library, January 1995

A:Description: A novel HIV-1 O strain illustrates the diversity of the O group.

A:Reference number: S52929

A:Accession: S52930

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-104 <COH>

A:Cross-references: EMBL:X84328; NID:g9695526; PIDN:CAA59066.1; PID:g695527

C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 159; DB 2; Length 104;

Best Local Similarity 81.2%; Pred. No. 3.4e-15;

Matches 26; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNIGCRGLVCYTSVRWNET 32

|||||:|||||:|||||:|||||:|||||

Db 35 ALETLMQNQLLNIGCRGLVCYTSVRWNET 66

RESULT 2

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborght, B.; van der Groen, (

J. Virol. 68, 1586-1596, 1994

A>Title: Genomic cloning and complete sequence analysis of a highly divergent African hu

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-863 <VAN>

A:Cross-references: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 88.1%; Score 155; DB 2; Length 863;

Best Local Similarity 81.2%; Pred. No. 9.4e-14;

Matches 26; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNIGCRGLVCYTSVRWNET 32

|||||:|||||:|||||:|||||:|||||

Db 584 ALETLLNQQLLNIGCRGLVCYTSVRWNET 615

```
RESULT 3
S49197
envelope protein precursor - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 26-Aug-1999
C;Accession: S49197
R;Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy,
submitted to the EMBL Data Library, July 1994
A;Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: defi
A;Reference number: S49197
A;Accession: S49197
A;Molecule type: DNA
A;Residues: 1-877 <CHA>
A;Cross-references: EMBL:X80020; NID:g510516; PIDN:CAA56323.1; PID:g510517
A;Experimental source: isolate VAU
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-535/Product: coat protein gp120 #status predicted <CP1>
F;536-877/Product: coat protein gp41 #status predicted <CP2>
F;698-716/Domain: transmembrane #status predicted <TMN>
F;57,86,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404,
Query Match 85.8%; Score 151; DB 2; Length 877;
Best Local Similarity 75.0%; Pred. No. 3.5e-13;
Matches 24; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWCGRLVCYTSVRNNE 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 597 ALETFIQQLLNINWCGKRLVCYTSVRNNE 628

RESULT 4
VCLJ51
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: simian immunodeficiency virus SIVcpz
A;Note: host Pan troglodytes (chimpanzee)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: S09990
R;Huet, T.; Cheynier, R.; Meyers, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A;Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A;Reference number: S09983; MUID:90259077; PMID:2188136
A;Accession: S09990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-854 <HUE>
A;Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36407.1; PID:g58874
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-500/Product: coat protein gp120 #status predicted <CP1>
F;501-854/Product: coat protein gp41 #status predicted <CP2>
F;501-517/Domain: transmembrane #status predicted <TM1>
F;675-693/Domain: transmembrane #status predicted <TM2>
F;805-821/Domain: transmembrane #status predicted <TM3>
F;134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,45
Query Match 64.8%; Score 114; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 5.4e-08;
Matches 18; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWCGRLVCYTSVRNNE 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 572 AVERYLDQQLINWCGSKAVCYTTPWNN 603

RESULT 5
env polyprotein D - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
```

```
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 20
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S21990; S70423
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
A;Reference number: S21990
A;Accession: S21990
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70423
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332, 'X', 334-357 <STE2>
A;Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C;Superfamily: type E retrovirus env polyprotein

Query Match 63.1%; Score 111; DB 2; Length 357;
Best Local Similarity 63.3%; Pred. No. 6.2e-08;
Matches 19; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWCGRLVCYTSVRWN 30
|||||:|||||:|||||:|||||:|||||:
Db 83 AVERYLDQQLINWCGSKRLICTTAVPWN 112

RESULT 6
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1, HIV-1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 3L
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S22002; S70418
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin
A;Reference number: S21990
A;Accession: S22002
A;Molecule type: DNA
A;Residues: 1-358 <STE1>
A;Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70418
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333, 'X', 335-358 <STE2>
A;Cross-references: EMBL:X61352; NID:g60186
C;Superfamily: type E retrovirus env polyprotein

Query Match 63.1%; Score 111; DB 2; Length 358;
Best Local Similarity 63.3%; Pred. No. 6.2e-08;
Matches 19; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALETLNQQLLNINWCGRLVCYTSVRWN 30
|||||:|||||:|||||:|||||:|||||:
Db 84 AVERYLDQQLINWCGSKRLICTTAVPWN 113

RESULT 7
B41621
env polyprotein D - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
```

N;Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 31-Jan-1997  
C;Accession: B41621  
R;Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991  
A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity  
A;Reference number: A41621; MUID:92107924; PMID:1763038  
A;Accession: B41621  
A;Molecule type: DNA  
A;Residues: 1-454 <BUR>  
A;Cross-references: GB:W72729  
A;Note: this virus was isolated from the daughter  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F;1-262/Product: coat protein gp120 (fragment) #status predicted <GP1>  
F;263-454/Product: coat protein gp41 (fragment) #status predicted <GP2>  
F;435-454/Domain: transmembrane #status predicted <TMN>  
F;9,23,36,42,48,49,78,85,102,108,132,138,142,195,210,214,362,367,376/Binding site: carboxyl  
Query Match 63.1%; Score 111; DB 2; Length 454;  
Best Local Similarity 63.3%; Pred. No. 7.8e-08;  
Matches 19; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 ALETLNQQLLNINWGCGRGLVCYTSVRWN 30  
Db 333 ALERYLKDQQLLGWCGSGKLICTTAVPWN 362  
RESULT 8  
S22000  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C;Accession: S22000  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by PCR  
A;Reference number: S21990  
A;Accession: S22000  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <STE>  
A;Cross-references: EMBL:X61351  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 62.5%; Score 110; DB 2; Length 358;  
Best Local Similarity 63.3%; Pred. No. 8.6e-08;  
Matches 19; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 ALETLNQQLLNINWGCGRGLVCYTSVRWN 30  
Db 84 AVERYLKDQQLLGWCGSGRLICTTAVPWN 113  
RESULT 9  
S70417  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 3B) (fragment)  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: patient 3B  
C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 26-Aug-1999  
C;Accession: S70417  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by PCR  
A;Reference number: S70417; MUID:92144209; PMID:1736940  
A;Accession: S70417  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <STE>

A;Cross-references: EMBL:X61351; NID:g60184; PIDN:CAA43614.1; PID:g60185  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 62.5%; Score 110; DB 2; Length 358;  
Best Local Similarity 63.3%; Pred. No. 8.6e-08;  
Matches 19; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 ALETLNQQLLNINWGCGRGLVCYTSVRWN 30  
Db 84 AVERYLKDQQLLGWCGSGRLICTTAVPWN 113  
RESULT 10  
C41621  
env polyprotein P - human immunodeficiency virus type 1 (fragment)  
N;Alternate names: coat polyprotein  
N;Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999  
C;Accession: C41621  
R;Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991  
A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity  
A;Reference number: A41621; MUID:92107924; PMID:1763038  
A;Accession: C41621  
A;Molecule type: DNA  
A;Residues: 1-443 <BUR>  
A;Cross-references: GB:W7230; NID:g328631; PIDN:AAB03792.1; PID:g555015  
A;Note: this virus was isolated from the mother's sexual partner  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F;1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>  
F;252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>  
F;424-443/Domain: transmembrane #status predicted <TMN>  
F;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: carboxyl  
Query Match 61.9%; Score 109; DB 2; Length 443;  
Best Local Similarity 56.2%; Pred. No. 1.5e-07;  
Matches 18; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 ALETLNQQLLNINWGCGRGLVCYTSVRWN 32  
Db 322 AVERYLKDQQLLGWCGSGKLICTTAVPWN 353  
RESULT 11  
S54384  
envelope polyprotein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Aug-1999  
C;Accession: S54384  
R;Theodore, T.; Buckler-White, A.J.  
submitted to the EMBL Data Library, July 1989  
A;Reference number: S54377  
A;Accession: S54384  
A;Status: preliminary  
A;Molecule type: genomic RNA  
A;Residues: 1-853 <THE>  
A;Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: polyprotein  
Query Match 61.9%; Score 109; DB 2; Length 853;  
Best Local Similarity 56.2%; Pred. No. 2.7e-07;  
Matches 18; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 ALETLNQQLLNINWGCGRGLVCYTSVRWN 32  
Db 579 AVERYLKDQQLLGWCGSGKLICTTAVPWN 610



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 8.06723 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362a-12  
Perfect score: 176  
Sequence: 1 ALETLQNOQLINIGCRGLVCTSVARNET 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	64.8	854	1 ENV_SIVCZ	P17281 chimpanzee
2	109	61.9	853	1 ENV_HV1Z2	P12487 human immun
3	109	61.9	855	1 ENV_HV1Z6	P04580 human immun
4	108	61.4	843	1 ENV_HV1Y2	P35961 human immun
5	108	61.4	847	1 ENV_HV1S1	P19550 human immun
6	108	61.4	847	1 ENV_HV1W2	P05880 human immun
7	108	61.4	851	1 ENV_HV1B8	P04582 human immun
8	108	61.4	852	1 ENV_HV1B3	P12488 human immun
9	108	61.4	852	1 ENV_HV1S3	P19549 human immun
10	108	61.4	853	1 ENV_HV1MF	P19551 human immun
11	108	61.4	855	1 ENV_HV1A2	P03378 human immun
12	108	61.4	855	1 ENV_HV1OY	P20688 human immun
13	108	61.4	856	1 ENV_HV1B1	P03375 human immun
14	108	61.4	856	1 ENV_HV1H2	P04578 human immun
15	108	61.4	856	1 ENV_HV1LW	O70626 human immun
16	108	61.4	856	1 ENV_HV1PV	P03376 human immun
17	108	61.4	856	1 ENV_HV1SC	P03678 human immun
18	108	61.4	856	1 ENV_HV1W1	P31872 human immun
19	108	61.4	861	1 ENV_HV1B1	P03377 human immun
20	108	61.4	865	1 ENV_HV1RH	P04579 human immun
21	108	61.4	867	1 ENV_HV1J3	P12489 human immun
22	106	60.2	848	1 ENV_HV1JR	P20871 human immun
23	106	60.2	856	1 ENV_HV1H3	P04624 human immun
24	106	60.2	856	1 ENV_HV1ZH	P05881 human immun
25	105	59.7	846	1 ENV_HV1ND	P18799 human immun
26	104	59.1	856	1 ENV_HV1MN	P05877 human immun
27	104	59.1	863	1 ENV_HV1Z8	P05882 human immun
28	104	59.1	868	1 ENV_HV1C4	P05881 human immun
29	102	58.0	853	1 ENV_HV1EL	P04581 human immun
30	102	58.0	861	1 ENV_HV1KB	P31819 human immun
31	96	54.5	859	1 ENV_HV1MA	P04583 human immun
32	89.5	50.9	856	1 ENV_HV2NZ	P05883 human immun
33	89.5	50.9	859	1 ENV_HV2D2	P15831 human immun

34	89.5	50.9	885	1 ENV_SIVS4	P12492 simian immu
35	88	50.6	857	1 ENV_HV2XR	O74126 human immu
36	88	50.0	858	1 ENV_HV2RO	P04577 human immu
37	87	49.4	821	1 ENV_SIVGB	P22380 simian immu
38	87	49.4	865	1 ENV_SIVAT	P05886 simian immu
39	86.5	49.1	712	1 ENV_HV2S2	P32536 human immu
40	86.5	49.1	859	1 ENV_HV2ST	P20872 human immu
41	86	48.9	877	1 ENV_SIVAG	P27977 simian immu
42	86	48.9	881	1 ENV_SIVMK	P05884 simian immu
43	86	48.9	882	1 ENV_SIVM1	P05885 simian immu
44	85.5	48.6	859	1 ENV_HV2CA	P24105 human immu
45	85	48.3	380	1 ENV_SIVM2	P08810 simian immu

## ALIGNMENTS

RESULT 1	ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;				
DT	01-AUG-1990 (Rel. 15, Created)				
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
GN	ENV.				
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_Taxid=11723;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90259077; PubMed=2188136;				
RA	Huet T., Cheylier R., Meyerhans A., Roelants G., Wain-Hobson S.,				
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."				
RL	Nature 345:356-359(1990).				
CC	-1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
CC	EMBL, X52154; CAA36407.1; -				
DR	PIR: S09990; VCLJSL.				
DR	HIV: X52154; ENVSCPZ.				
DR	InterPro: IPR000328; Env GP41.				
DR	InterPro: IPR000777; GP120.				
DR	Pfam: PF00516; GP120, 1.				
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;				
FT	SIGNAL	1	30		
FT	CHAIN	31	500		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854		TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517		POTENTIAL.
FT	TRANSMEM	675	693		POTENTIAL.
FT	TRANSMEM	805	821		POTENTIAL.
FT	CARBOHYD	134	134		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	140	140		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	143	143		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	154	154		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	158	158		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	186	186		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	195	195		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	239	239		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	260	260		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	267	267		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	274	274		N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	299	299		N-LINKED (GLCNAC. . .) (POTENTIAL)

```

FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 64.8%; Score 114; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 2.6e-09;
Matches 18; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

1 ALETLGNQOLLNMGCRGLVCYTSVNMNET 32
Db 572 AVERYLDDQOLLGIMGSGKAVCITTVPMNNS 603

RESULT 2
ENV_HV122 STANDARD; PRT; 853 AA.
ID _ENV_HV122
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (Z2/CDC-Z24 isolate) (HIV-1).
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T. Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV$2226.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

Query Match 61.9%; Score 109; DB 1; Length 853;
Best Local Similarity 56.2%; Pred. No. 1.5e-08;
Matches 18; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

1 ALETLGNQOLLNMGCRGLVCYTSVNMNET 32
Db 579 AVERYLDDQOLLGIMGSGKLCITTVPMNNS 610

RESULT 3
ENV_HV126 STANDARD; PRT; 855 AA.
ID _ENV_HV126
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; Pubmed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaraman V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.

```

CC	Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC	Vitruvisee; Retroviridae; Lentivirus.
OK	NCBI TaxID=36377;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=93021387; PubMed=1404605;
RA	Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H., Shaw G.M.;
RT	"Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation.";
RT	J. Virol. 66:6587-6600(1992).
RL	
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way used modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC	
DR	EMBL; M93258; ; NOT_ANNOTATED_CDS.
DR	PDB; H44001; H44001.
DR	PDB; 1G9N; 27-DEC-00.
DR	InterPro; IPR000328; Env_GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal; 3D-structure.
KW	SIGNAL
FT	CHAIN 1 29
FT	CHAIN 30 489
FT	TRANSMEM 490 843
FT	TRANSMEM 738 755
FT	DISEULFID 53 73
FT	DISEULFID 118 201
FT	DISEULFID 125 192
FT	DISEULFID 130 155
FT	DISEULFID 214 243
FT	DISEULFID 224 235
FT	DISEULFID 292 326
FT	DISEULFID 373 432
FT	DISEULFID 380 405
FT	CARBOHYD 87
FT	CARBOHYD 129 129
FT	CARBOHYD 135 135
FT	CARBOHYD 138 138
FT	CARBOHYD 154 154
FT	CARBOHYD 158 158
FT	CARBOHYD 184 184
FT	CARBOHYD 193 193
FT	CARBOHYD 230 230
FT	CARBOHYD 237 237
FT	CARBOHYD 258 258
FT	CARBOHYD 272 272
FT	CARBOHYD 285 285
FT	CARBOHYD 291 291
FT	CARBOHYD 297 297
FT	CARBOHYD 327 327
FT	CARBOHYD 351 351
FT	CARBOHYD 381 381
FT	CARBOHYD 389 389
FT	CARBOHYD 395 395
FT	CARBOHYD 400 400
FT	CARBOHYD 435 435
FT	CARBOHYD 450 450
FT	CARBOHYD 598 598
FT	CARBOHYD 603 603
FT	CARBOHYD 612 612
FT	CARBOHYD 624 624
FT	CARBOHYD 803 803
QO	SEQUENCE 843 AA; 95648 MW; C69DFD971C91B871 CRC64;



Query Match 61.4%; Score 108; DB 1; Length 843;  
 Best Local Similarity 60.0%; Pred. No. 2,1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALETTLQNOQLINWGRGLVCTSVRN 30  
 569 AVERLYRDKOQLLGWCGSKLCTTAVPMN 598

RESULT 5  
 ENV\_HV1S1 STANDARD; PRT; 847 AA.

AC P19550;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retrovirinae; Lentivirus.  
 OX NCBI\_TaxID=11691;  
 RN [1]  
 RP MEDLINE=90347835; PubMed=2384920;  
 RA Cheng-Mayer C., Quiroga M., Tung J.W., Levy J.,  
 RT "Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";  
 RL J. Virol. 64:4390-4398(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M65024; AAA5072.1; -.  
 DR PDB; 1OBR; 15-MAY-97.  
 DR HTV; M38428; ENV\$SF162.  
 DR InterPro; IPR000328; Env GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KM AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal; 3D-structure.  
 KW SIGNAL 1 29  
 FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 203 BY SIMILARITY.  
 FT DISULFID 125 194 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 216 245 BY SIMILARITY.  
 FT DISULFID 226 237 BY SIMILARITY.  
 FT DISULFID 294 328 BY SIMILARITY.  
 FT DISULFID 374 435 BY SIMILARITY.  
 FT DISULFID 381 408 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7F2AB CRC64;

Query Match 61.4%; Score 108; DB 1; Length 847;  
 Best Local Similarity 60.0%; Pred. No. 2,1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALETTLQNOQLINWGRGLVCTSVRN 30  
 573 AVERLYRDKOQLLGWCGSKLCTTAVPMN 602

RESULT 6  
 ENV\_HV1W2 STANDARD; PRT; 847 AA.

AC P05880;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retrovirinae; Lentivirus.  
 OX NCBI\_TaxID=11705;  
 RN [1]  
 RP MEDLINE=86235450; PubMed=1012778;  
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
 RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.,  
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS.";  
 RL Science 232:1548-1553(1986).  
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M12507; AAB12990.1; -.  
 DR HTV; M12507; ENV\$WMJ2.  
 DR InterPro; IPR000328; Env GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KM AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal.  
 KW SIGNAL 1 29  
 FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 202 BY SIMILARITY.  
 FT DISULFID 125 193 BY SIMILARITY.  
 FT DISULFID 130 152 BY SIMILARITY.  
 FT DISULFID 215 244 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 293 326 BY SIMILARITY.



```
FT DISULFID 372 435 BY SIMILARITY.
FT DISULFID 379 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96466 MW; CDIE33D73AA5BCAE CRC64;
```

```
Query Match 61.4%; Score 108; DB 1; Length 847;
Best Local Similarity 60.0%; Pred. No. 2.1e-08;
Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy 1 ALETLNQQLINWGCRGLVCTSVRN 30
Db 573 AVERLRDQQLIGWGSGKLCTTVPWN 602
```

```
RESULT 7
ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (B8 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Liyak K.J., Starch B.R.,
RA Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumbarger K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Laubenstein J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284 (1985).
```

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [licences@isb-sib.ch](mailto:licences@isb-sib.ch)).  
-----

```
DR EMBL; K02011; AAA44661.1; -
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HG3; 31-OCT-93.
DR PDB; 1Q03; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENVSH8.
DR GlycoSuiteDB; P04582; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 506
FT DISULFID 507 851 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.
FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;
```

```
Query Match 61.4%; Score 108; DB 1; Length 851;
Best Local Similarity 60.0%; Pred. No. 2.1e-08;
Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy 1 ALETLNQQLINWGCRGLVCTSVRN 30
Db 577 AVERLRDQQLIGWGSGKLCTTVPWN 606
```

```
RESULT 8
ENV_HV1B8 STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
```

DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 NC NCBI\_TaxID=11693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=69085613; PubMed=2789516;  
 RA Andard R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,  
 RT Dandekar S.,  
 RT "Biological and molecular characterization of human immunodeficiency  
 RT virus (HIV-1BR) from the brain of a patient with progressive  
 RT dementia.";  
 RL Virology 168:79-89 (1989).  
 CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS  
 CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M21098; AAA44221.1; -  
 CC PIR; A31667; VCLJBR.  
 CC PDB; 1IM7; 23-OCT-02.  
 CC HIV; M21098; ENV5BRVA.  
 CC InterPro; IPR000328; Env\_GP41.  
 CC InterPro; IPR000777; GP120.  
 CC Pfam; PF00516; GP120; 1.  
 CC Pfam; PF00517; GP41; 1.  
 CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 CC 3D-structure.  
 CC KW  
 CC SIGNAL 1 30  
 CC CHAIN 31 507  
 CC FT CHAIN 508 852  
 CC FT DISULFID 54 74  
 CC FT DISULFID 119 205  
 CC FT DISULFID 126 196  
 CC FT DISULFID 131 155  
 CC FT DISULFID 218 247  
 CC FT DISULFID 228 239  
 CC FT DISULFID 296 330  
 CC FT DISULFID 376 439  
 CC FT DISULFID 383 412  
 CC FT CARBOHYD 49 49  
 CC FT CARBOHYD 88 88  
 CC FT CARBOHYD 135 135  
 CC FT CARBOHYD 138 138  
 CC FT CARBOHYD 154 154  
 CC FT CARBOHYD 158 158  
 CC FT CARBOHYD 197 197  
 CC FT CARBOHYD 234 234  
 CC FT CARBOHYD 241 241  
 CC FT CARBOHYD 262 262  
 CC FT CARBOHYD 276 276  
 CC FT CARBOHYD 289 289  
 CC FT CARBOHYD 295 295  
 CC FT CARBOHYD 301 301  
 CC FT CARBOHYD 331 331  
 CC FT CARBOHYD 354 354  
 CC FT CARBOHYD 360 360  
 CC FT CARBOHYD 384 384  
 CC FT CARBOHYD 390 390  
 CC FT CARBOHYD 396 396  
 CC FT CARBOHYD 400 400  
 CC FT CARBOHYD 442 442  
 CC FT CARBOHYD 456 456  
 CC FT CARBOHYD 607 607  
 CC FT CARBOHYD 612 612  
 CC FT CARBOHYD 621 621  
 CC FT CARBOHYD 621 621

FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;  
 Query Match 61.4%; Score 108; DB 1; Length 852;  
 Best Local Similarity 60.0%; Pred. No. 2.1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 ALETLQNOQLNIMGCRGRVCTSVRM 30  
 Db 578 AVERYLKQQLGIMGSGKLCITTAAPWN 607  
 RESULT 9  
 ENV\_HV1S3 STANDARD; PRT; 852 AA.  
 AC P19549;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317906; PubMed=2370688;  
 RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.,  
 RT "Human immunodeficiency virus type 1 cellular host range,  
 RT replication, and cytopathicity are linked to the envelope region of  
 RT the viral genome.";  
 RL J. Virol. 64:4016-4020(1990).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AY52275; AAQ17031.1; -  
 CC PDB; 1MEQ; 11-DEC-02.  
 CC HIV; M38427; ENV5SRF3.  
 CC InterPro; IPR000328; Env\_GP41.  
 CC InterPro; IPR000777; GP120.  
 CC Pfam; PF00516; GP120; 1.  
 CC Pfam; PF00517; GP41; 1.  
 CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 CC 3D-structure.  
 CC KW  
 CC SIGNAL 1 31  
 CC CHAIN 32 506  
 CC FT CHAIN 507 852  
 CC FT DISULFID 53 73  
 CC FT DISULFID 118 206  
 CC FT DISULFID 125 197  
 CC FT DISULFID 130 156  
 CC FT DISULFID 219 248  
 CC FT DISULFID 229 240  
 CC FT DISULFID 297 331  
 CC FT DISULFID 377 439  
 CC FT DISULFID 384 412  
 CC FT CARBOHYD 87 87  
 CC FT CARBOHYD 129 129  
 CC FT CARBOHYD 136 136  
 CC FT CARBOHYD 141 141  
 CC FT CARBOHYD 142 142  
 CC FT CARBOHYD 155 155  
 CC FT CARBOHYD 159 159  
 CC FT CARBOHYD 189 189  
 CC FT CARBOHYD 189 189

```

FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 852 AA; 96663 MW; E87BFBFD3C9910D CRC64;

Query Match 61.4%; Score 108; DB 1; Length 852;
Best Local Similarity 60.0%; Pred. No. 2.1e-08;
Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Cy 1 ALETLNQQLNTWGCGRGLVCTSVYRN 30
Db 578 AVERLYKQQLLGIGSGKLCITTVPMN 607

RESULT 10
ENV_HV1MP STANDARD; PRT; 853 AA.
ID ENV_HV1MP
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonic C., Mann A.M., Meier C.,
RA Wasiak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytolysis."
RT J. Virol. 64:3792-3803(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M33943; AAA44850.1; -
CC PDB; 1A1K; 16-JUN-97.
CC HIV; M33943; ENVSMFA.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120.1.
CC Pfam; PF00517; GP41.1.
CC AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
CC 3D-structure. 1 30
FT SIGNAL

```

```

FT CHAIN 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 853 AA; 96912 MW; 337B93B6F22ABA CRC64;

Query Match 61.4%; Score 108; DB 1; Length 853;
Best Local Similarity 60.0%; Pred. No. 2.1e-08;
Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Cy 1 ALETLNQQLNTWGCGRGLVCTSVYRN 30
Db 580 AVERLYKQQLLGIGSGKLCITTVPMN 609

RESULT 11
ENV_HV1A2 STANDARD; PRT; 855 AA.
ID ENV_HV1A2
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stemple M.W., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2)."
RT Science 227:484-492(1985).
FT SIGNAL

```



FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607AD62DA CRC64;  
 Query Match 61.4%; Score 108; DB 1; Length 855;  
 Best Local Similarity 60.0%; Pred. No. 2.1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 ALETLNQOQLNMGCRGLVCTSVRN 30  
 DB 581 AVERLYKQQLGIGSGKLICTTVPWN 610  
 RESULT 13  
 ENV\_HV1B1 STANDARD; PRT; 856 AA.  
 ID ENV\_HV1B1  
 AC P03375;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8511123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Lyvak K.J., Searl B.R.,  
 Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 Baumanster K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
 Lautenberger J.A., Pagas T.S., Ghayab J., Chang N.T., Gallo R.C.,  
 Wong-Staal F.;  
 RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RT Nature 313:277-284 (1985).  
 RL J. Biol. Chem. 265:10373-10382 (1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; M15654; AAA44205.1; -  
 DR PIR; A03973; VCLJH3.  
 DR HIV; M15654; ENV5BHL02.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157

FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .)  
 SQ SEQUENCE 856 AA; 97224 MW; 0BFB1A1931BB27 CRC64;  
 Query Match 61.4%; Score 108; DB 1; Length 856;  
 Best Local Similarity 60.0%; Pred. No. 2.1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 ALETLNQOQLNMGCRGLVCTSVRN 30  
 DB 582 AVERLYKQQLGIGSGKLICTTAVPWN 611  
 RESULT 14  
 ENV\_HV1B2 STANDARD; PRT; 856 AA.  
 ID ENV\_HV1B2  
 AC P04578; 009779;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11706;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87299196; PubMed=3040055;  
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
 Gallo R.C., Wong-Staal F.;  
 RA "Complete nucleotide sequences of functional clones of the AIDS  
 RT virus.";  
 RT AIDS Res. Hum. Retroviruses 3:57-69 (1987).  
 RL [2]  
 RN REVISIONS.  
 RP Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
 Gallo R.C., Wong-Staal F.;  
 RA Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.



FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA;	96938 MW;	0C241332CF7E6687	CRC64;

Query Match 61.4%; Score 108; DB 1; Length 856;  
 Best Local Similarity 60.0%; Pred. No. 2, 1e-08;  
 Matches 18; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 1 ALETLNQQLNINWCGRGRLVVCYTSVRWN 30  
 DB 582 AVERYLKQDLGIWCGSKLICTTAVPWN 611

Search completed: May 7, 2004, 17:43:56  
 Job time : 8.06723 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 40.605 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-12

Perfect score: 176

Sequence: 1 ALETLLNQQLLNIGCRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	165	93.8	124	15	Q9IHU7 human immun
2	165	93.8	126	15	Q9IHV1 human immun
3	165	93.8	172	15	Q9IEB3 human immun
4	165	93.8	234	15	Q9IEC2 human immun
5	164	93.2	240	15	Q9IE32 human immun
6	163	92.6	125	15	Q9IHU8 human immun
7	163	92.6	216	15	Q9IEA5 human immun
8	163	92.6	544	15	Q9IED9 human immun
9	162	92.0	116	15	O40459 human immun
10	162	92.0	137	15	Q9IHV5 human immun
11	162	92.0	155	15	Q8J3Q4 human immun
12	162	92.0	208	15	Q9IEA3 human immun
13	162	92.0	216	15	Q9IEC5 human immun
14	162	92.0	238	15	Q9DIK1 human immun
15	162	92.0	342	15	O11942 human immun
16	162	92.0	418	15	O36547 human immun

17	162	92.0	871	15	O57074	057074 human immun
18	162	92.0	871	15	O8Q7I2	O8Q7I2 human immun
19	162	92.0	876	15	O8Q7H3	O8Q7H3 human immun
20	161	91.5	105	15	O11939	O11939 human immun
21	161	91.5	114	15	O40472	O40472 human immun
22	161	91.5	118	15	O40451	O40451 human immun
23	161	91.5	213	15	Q9IEC4	Q9IEC4 human immun
24	161	91.5	532	15	Q9IEF0	Q9IEF0 human immun
25	161	91.5	872	15	Q900Y5	Q900Y5 human immun
26	161	91.5	880	15	O8Q7H1	O8Q7H1 human immun
27	160	90.9	116	15	O40458	O40458 human immun
28	160	90.9	134	15	Q9IHV4	Q9IHV4 human immun
29	160	90.9	158	15	O8J3N6	O8J3N6 human immun
30	160	90.9	183	15	Q9IEC1	Q9IEC1 human immun
31	160	90.9	209	15	Q9IE66	Q9IE66 human immun
32	160	90.9	220	15	Q9IEC9	Q9IEC9 human immun
33	160	90.9	225	15	Q9IEA0	Q9IEA0 human immun
34	160	90.9	512	15	Q9IED2	Q9IED2 human immun
35	160	90.9	545	15	Q9IED7	Q9IED7 human immun
36	160	90.9	875	15	O8Q7H4	O8Q7H4 human immun
37	160	90.9	879	15	Q9WIU9	Q9WIU9 human immun
38	160	90.9	900	15	Q9QNZ8	Q9QNZ8 human immun
39	159	90.3	104	15	O76163	O76163 human immun
40	159	90.3	114	15	O40457	O40457 human immun
41	159	90.3	116	15	Q7ZJN9	Q7ZJN9 human immun
42	159	90.3	200	15	Q9IEB8	Q9IEB8 human immun
43	159	90.3	224	15	Q9IEA8	Q9IEA8 human immun
44	159	90.3	230	15	Q9IEB2	Q9IEB2 human immun
45	159	90.3	517	15	Q9IEE7	Q9IEE7 human immun

ALIGNMENTS

RESULT 1

Q9IHU7 ID Q9IHU7 PRELIMINARY; PRT; 124 AA.  
AC Q9IHU7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97CM768;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,  
RA Pieniazek D., Schable C., Lal R.B.;  
RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O.";  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
DR EMBL; AF229237; AAF71914.1; -  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 15187 MW; 105D515F114450F8 CRC64;

Query Match 93.8%; Score 165; DB 15; Length 124;  
Best Local Similarity 84.4%; Pred. No. 4.3e-17;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLNIGCRGLVCYTSVRWNET 32

Db 23 ALETLLNQQLLNIGCRGLVCYTSVRWNET 54

Query Match	93.8%	Score 165	DB 15	Length 172
Best Local Similarity	84.4%	Pred. NO. 6e-17		
Matches	27	Conservative	5	Mismatches 0
				Indels 0
				Gaps 0
QY	1	ALETLLQOOLLNIWCGRGLVCYTSVRWNET	32	
DB	26	ALETLLQOOLLNIWCGKGRGLCYTSVRWNT	57	
<p>     :     :     :     :     : </p>				
<p>RESULT 4</p>				
Q9IEC2	Q9IEC2	PRELIMINARY;	PRT;	234 AA.
ID	Q9IEC2	PRELIMINARY;	PRT;	234 AA.
AC	Q9IEC2	AC		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	GP41	(Fragment).		
GN	ENV.			
OS	Human immunodeficiency virus 1.			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11676;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BCF100;			
RA	Roques P., Robertson D., Sandrine S., Christel D., Francois S.,			
RA	Philippe M.;			
RT	"Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";			
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; A236397; CAB96246.1; --			
DR	GO; GO:0016021; C:integral to membrane; IEA.			
DR	GO; GO:0019031; C:viral envelope; IEA.			
DR	GO; GO:0005198; F:structural molecule activity; IEA.			
DR	InterPro; IPR000328; Env_GP41.			
DR	Pfam; PF00517; GP41; 1.			
KW	Transmembrane.			
FT	NON TER 1			
FT	NON TER 234			
FT	NON TER 234			
SEQ	SEQUENCE 234 AA; 27036 MW; 39B050B3F8555A8C CRC64;			
<p>Query Match 93.8%; Score 165; DB 15; Length 234;</p> <p>Best Local Similarity 84.4%; Pred. NO. 8.1e-17;</p> <p>Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;</p>				
QY	1	ALETLLQOOLLNIWCGRGLVCYTSVRWNET	32	
DB	48	ALETLLQOOLLNIWCGKGRGLCYTSVRWNT	79	
<p>     :     :     :     :     : </p>				
<p>RESULT 5</p>				
Q9IE32	Q9IE32	PRELIMINARY;	PRT;	240 AA.
ID	Q9IE32	PRELIMINARY;	PRT;	240 AA.
AC	Q9IE32	AC		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	TM, GP41	(Fragment).		
GN	ENV.			
OS	Human immunodeficiency virus 1.			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11676;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BCF111;			
RA	Roques P., Robertson D., Sousquiere S., Damond F., Mauchere P.,			
RA	Delienne C., Brun-Vesinet F., Dormont D., Simon F.O.;			
RT	"Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";			
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; A2343365; CAB96335.1; --			
DR	GO; GO:0016021; C:integral to membrane; IEA.			
DR	GO; GO:0019031; C:viral envelope; IEA.			
DR	GO; GO:0005198; F:structural molecule activity; IEA.			
DR	InterPro; IPR000328; Env_GP41.			
DR	Pfam; PF00517; GP41; 1.			

```

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236415; CAB96263.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR DR GO; GO:0005198; F:structural molecule activity; IEA.
DR DR InterPro; IPR000328; Env_GP41.
DR DR Pfam; PF00517; GP41; 1.
DR DR Transmembrane.
KW KW
FT FT NON_TER 1 1
FT FT NON_TER 216 216
SQ SEQUENCE 216 AA; 25003 MW; 0A5AC218BFA88932 CRC64;

Query Match 92.6%; Score 163; DB 15; Length 216;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLQNOLLNMGCRGLVCYTSVRWNNT 32
Db 32 ALETLQNOLLNMGCKGRGLVCYTSVRWNNT 63
|||||:|||||:|||||:|||||:|||||
|||||:|||||:|||||:|||||:|||||

RESULT 8
O91ED9 PRELIMINARY; PRT; 544 AA.
ID O91ED9 AC O91ED9 AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
DE ENV.
GN Human immunodeficiency virus 1.
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
OX [1] RN
XP SEQUENCE FROM N.A.
RP STRAIN=YB26;
RC
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RT Depienne C., Brun-Vezinet F., Dormont D.;
RL "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13068; CAB96229.1; -.
DR DR GO; GO:0016021; C:integral to membrane; IEA.
DR DR GO; GO:0019028; C:viral capsid; IEA.
DR DR GO; GO:0019031; C:viral envelope; IEA.
DR DR GO; GO:0005198; F:structural molecule activity; IEA.
DR DR InterPro; IPR000328; Env_GP41.
DR DR InterPro; IPR000777; GP120.
DR DR Pfam; PF00516; GP120; 1.
DR DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT FT NON_TER 1 1
FT FT NON_TER 544 544
SQ SEQUENCE 544 AA; 61398 MW; FC3CFAE31DB6D50 CRC64;

Query Match 92.6%; Score 163; DB 15; Length 544;
Best Local Similarity 87.5%; Pred. No. 3.7e-16;
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLQNOLLNMGCRGLVCYTSVRWNNT 32
Db 360 ALETLQNOLLNMGCKGRGLVCYTSVRWNNT 391
|||||:|||||:|||||:|||||:|||||
|||||:|||||:|||||:|||||:|||||

RESULT 9
O40459 PRELIMINARY; PRT; 116 AA.
ID O40459 AC O40459 AC
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope transmembrane glycoprotein (Fragment).
DE Human immunodeficiency virus 1.
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```

OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=group O;
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;
RT "Molecular characterization of envelope transmembrane glycoprotein of
RT 14 new human immunodeficiency virus type 1 group O strains from
RT different African countries.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09775; CAA70914.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13975 MW; 12B3D0D0D2A1AD32 CRC64;

Query Match          92.0%; Score 162; DB 15; Length 116;
Best Local Similarity 84.4%; Pred. No. 1.2e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWCGRLVCYTSVRWNET 32
Db 29 ALETLNQQLLNWCGKRLVCYTSVKWNT 60

RESULT 10
Q9IHV5 Q9IHV5 PRELIMINARY; PRT; 137 AA.
AC Q9IHV5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM761;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,
RA Pléniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081 (2000).
DR EMBL; AF229229; AAF71906.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 16494 MW; 55C47096D8168493 CRC64;

Query Match          92.0%; Score 162; DB 15; Length 137;
Best Local Similarity 84.4%; Pred. No. 1.4e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWCGRLVCYTSVRWNET 32
Db 30 ALETLNQQLLNWCGKRLVCYTSVKWNT 61

RESULT 11
Q8J3Q4 Q8J3Q4 PRELIMINARY; PRT; 155 AA.
ID Q8J3Q4
```

```

AC Q8J3Q4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 protein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22092513; PubMed=12097573;
RA Boutonnet N., Janssens W., Boutton C., Verschelde J.L., Heyndrickx L.,
RA Beirnaert E., van der Groen G., Lasters I.;
RT "Comparison of predicted scaffold-compatible sequence variation in the
RT triple-hairpin structure of human immunodeficiency virus type 1 gp41
RT with patient data.";
RL J. Virol. 76:7595-7606 (2002).
DR EMBL; AJ427999; CAD20951.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 155 155
FT NON_TER 155 AA; 18571 MW; 1EA7198FFFP9BB1F CRC64;

Query Match          92.0%; Score 162; DB 15; Length 155;
Best Local Similarity 81.2%; Pred. No. 1.5e-16;
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWCGRLVCYTSVRWNET 32
Db 38 ALETLNQQLLNWCGKRLVCYTSVKWNT 69

RESULT 12
Q9IEA3 Q9IEA3 PRELIMINARY; PRT; 208 AA.
ID Q9IEA3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF32;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236417; CAB96265.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 24286 MW; 1C2961C1953A07A1 CRC64;

Query Match          92.0%; Score 162; DB 15; Length 208;
Best Local Similarity 81.2%; Pred. No. 2.1e-16;
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLNWCGRLVCYTSVRWNET 32
Db 1 ALETLNQQLLNWCGKRLVCYTSVKWNT 32
```

Db 24 ALETLLNQQLNLWGCKGLICYTSVKWNET 55

## RESULT 13

Q9IEC5 Q9IEC5 PRELIMINARY; PRT; 216 AA.  
AC Q9IEC5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF07;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ236394; CAB96243.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0019031; C: viral envelope; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
DR Transmembrane. 1  
FT NON\_TER 1  
FT NON\_TER 216  
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 92.0%; Score 162; DB 15; Length 216;  
Best Local Similarity 87.5%; Pred. No. 2.1e-16;  
Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNLWGCKGLICYTSVKWNET 32  
Db 37 ALETLLNQQLNLWGCKGLICYTSVKWNET 68

## RESULT 14

Q9DIK1 Q9DIK1 PRELIMINARY; PRT; 238 AA.  
AC Q9DIK1;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF119;  
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,  
Depienne C., Brun-Vezinet F., Dormont D., Simon F.;  
RT "Phylogenetic analysis of 49 newly derived HIV-1 group O strains: High  
viral diversity but no group M-like subtypes."  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ298127; CAC18808.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0019031; C: viral envelope; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
DR Transmembrane. 1  
FT NON\_TER 1  
FT NON\_TER 238  
SQ SEQUENCE 238 AA; 27585 MW; C5164D3B177BDA78 CRC64;

Query Match 92.0%; Score 162; DB 15; Length 238;  
Best Local Similarity 84.4%; Pred. No. 2.4e-16;  
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNLWGCKGLICYTSVKWNET 32  
Db 52 ALETLLNQQLNLWGCKGLICYTSVKWNET 83

## RESULT 15

O11942 O11942 PRELIMINARY; PRT; 342 AA.  
AC O11942;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ABT2156;  
RX MEDLINE=97340911; PubMed=9197385;  
RA Brennan C.A., Hackett J. Jr., Zekeng L., Lund J.K., Vallari A.S.,  
Hickman R.K., Gurtler L., Kaptue L., von Overbeck J., Hampl H.,  
Devare S.G.;  
RT "Sequence of gp41env immunodominant region of HIV type 1 group O from  
west central Africa."  
RL AIDS Res. Hum. Retroviruses 13:901-904 (1997).  
DR EMBL: U90135; AAB62818.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0019031; C: viral envelope; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
DR Transmembrane. 1  
FT NON\_TER 1  
FT NON\_TER 342  
SQ SEQUENCE 342 AA; 38464 MW; E337FEC8A29B67A1 CRC64;

Query Match 92.0%; Score 162; DB 15; Length 342;  
Best Local Similarity 84.4%; Pred. No. 3.4e-16;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLNLWGCKGLICYTSVKWNET 32  
Db 65 ALETLLNQQLNLWGCKGLICYTSVKWNET 96

Search completed: May 7, 2004, 17:49:14  
Job time : 41.605 secs

THIS PAGE BLANK (cont)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	176	100.0	32	2	AAW80471	AAW80471 Peptide d
2	169	96.0	32	2	AAW80470	AAW80470 Peptide d
3	165	93.8	32	2	AAW80469	AAW80469 Peptide d
4	162	92.0	200	3	AAW77373	AAW77373 HIV-1 gro
5	162	92.0	215	2	AAW09499	AAW09499 HIV-1 Gro
6	162	92.0	215	2	AAW06983	AAW06983 Recombina
7	162	92.0	215	3	AAW77374	AAW77374 HIV-1 gro
8	162	92.0	245	2	AAW09493	AAW09493 HIV-1 Gro
9	162	92.0	245	2	AAW06977	AAW06977 Recombina
10	162	92.0	245	3	AAW77369	AAW77369 HIV-1 gro
11	162	92.0	281	2	AAW09507	AAW09507 HIV-1 Gro
12	162	92.0	373	2	AAW09495	AAW09495 HIV-1 Gro
13	162	92.0	373	2	AAW06979	AAW06979 Recombina
14	162	92.0	460	2	AAW09500	AAW09500 HIV-1 Gro
15	162	92.0	460	2	AAW06984	AAW06984 Recombina
16	162	92.0	460	3	AAW77375	AAW77375 HIV-1 gro
17	162	92.0	474	3	AAW77371	AAW77371 HIV-1 gro
18	162	92.0	488	2	AAW09504	AAW09504 HIV-1 Gro
19	162	92.0	490	2	AAW09494	AAW09494 HIV-1 Gro
20	162	92.0	490	2	AAW06978	AAW06978 Recombina
21	162	92.0	490	3	AAW77370	AAW77370 HIV-1 gro
22	162	92.0	526	2	AAW09505	AAW09505 HIV-1 Gro
23	162	92.0	618	2	AAW09496	AAW09496 HIV-1 Gro
24	162	92.0	618	2	AAW06980	AAW06980 Recombina
25	162	92.0	618	3	AAW77372	AAW77372 HIV-1 gro

```

Best Local Similarity 100.0%; Pred. No. 1.6e-15;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCRGLVCYTSVRWNET 32
    |||||
Db 1 ALETLNQQLLDLWGCRGLVCYTSVRWNET 32

RESULT 2
AAW80470
ID AAW80470 standard; peptide; 32 AA.
XX
AC AAW80470;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
XX Peptide derived from a conserved sequence of group O human HIV.
DE Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
OS Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
PA
XX
XX Chenebaux DMB, Delagneau JH, Gabelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
XX Claim 6; Page 44; 55pp; French.
XX
XX CAAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 32 AA;

Query Match 96.0%; Score 169; DB 2; Length 32;
Best Local Similarity 93.8%; Pred. No. 1.3e-14;
Matches 30; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCRGLVCYTSVRWNET 32
    |||||
Db 1 ALETLNQQLLDLWGCRGLVCYTSVRWNET 32

RESULT 3
AAW80469
ID AAW80469 standard; peptide; 32 AA.
XX
XX AAW80469;
AC
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
DT

```







```

XX PD 25-FEB-1999.
XX PF 17-AUG-1998; 98WO-US017014.
XX PR 15-AUG-1997; 97US-00911824.
XX PA (ABBO ) ABBOTT LAB.
XX PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
XX DR WPI; 1999-190167/16.
XX DR N-PSDB; AAX56074.
XX PT New isolated HIV-1 Group O env polypeptides - used for the detection of
XX PT anti-HIV antibodies and for the production of antibodies for use in
XX PT detection, purification and therapy.
XX PS Claim 16; Fig 7; 138pp; English.
XX
CC The present invention describes (A) an isolated HIV-1 Group O env
CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env
CC polypeptide comprising an immunoreactive portion of a polypeptide as in
CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at
CC least one HIV-1 Group M env polypeptide; (5) an antigen construct
CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an
CC antigen construct comprising a first HIV-2 env polypeptide fused to a
CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
CC host cell transformed by an expression vector as in (8); and (10) an
CC immunoassay kit for the detection of antibodies to HIV-1 comprising an
CC antigen construct as in (3)-(6). The antigen constructs can be used for
CC the detection of anti-HIV-1 antibodies in test samples. They can also be
CC used as immunogens to produce antibodies. The antibodies can be used to
CC purify HIV polypeptides, for therapy and for detection of HIV
CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 245 AA;
Query Match 92.0%; Score 162; DB 2; Length 245;
Best Local Similarity 84.4%; Pred. No. 7.6e-13;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 32
Db 117 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 148
RESULT 9
AAY06977
ID AAY06977 standard; protein; 245-AA.
XX
AC AAY06977;
XX
DT 06-JUL-1999 (first entry)
XX
DE Recombinant pGO-9PL protein.
XX
KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;
XX antibody; assay.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
XX
FH Key Location/Qualifiers
FT Protein 2..46
FT Peptide /note= "gp120 sequence"
FT 47..245
FT /note= "gp41 sequence"

```

```

XX WO9909410-A2.
XX PN 25-FEB-1999.
XX PD
XX PF 07-AUG-1998; 98WO-US016506.
XX PR 15-AUG-1997; 97US-00912129.
XX PA (ABBO ) ABBOTT LAB.
XX PI Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
XX PI Golden AM, Brennan CA, Devare SG;
XX DR WPI; 1999-190224/16.
XX DR N-PSDB; AAX37189.
XX PT New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
XX PT be used in field assay, requiring no electricity and less specialised
XX PT equipment.
XX PS Claim 1; Fig 7; 104pp; English.
XX
CC The invention relates to a rapid assay for simultaneous detection and
CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
CC method comprises (a) contacting the sample with a strip containing at
CC least one immobilised capture reagent per analyte and on which the sample
CC moves from the proximal to the distal end by capillary action, under
CC conditions sufficient to form capture reagent/analyte complexes, and (b)
CC determining the presence of analyte(s) by detecting a visible colour
CC change at the capture reagent site on the strip wherein the capture
CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in
CC AAY06981. The invention is used to screen patients for antibodies to HIV-
CC 1 types O and M, and HIV-2. The invention will be particularly useful in
CC places and situation where equipment and/or electricity is not available.
CC The invention provides a screening method which is faster and requires
CC less equipment than prior art methods. The present sequence represents a
CC amino acid sequence of the recombinant pGO-9PL recombinant protein which
CC acts as a capture reagent for HIV-1 group O
XX
SQ Sequence 245 AA;
Query Match 92.0%; Score 162; DB 2; Length 245;
Best Local Similarity 84.4%; Pred. No. 7.6e-13;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 32
Db 117 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 148
RESULT 10
AAY77369
ID AAY77369 standard; protein; 245-AA.
XX
AC AAY77369;
XX
DT 22-MAY-2000 (first entry)
XX
DE HIV-1 group O pGO-9PL-encoded truncated env gp41 protein, SEQ ID NO:48.
XX
KW HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;
XX immunoassay; positive control; affinity purification; therapeutic;
XX Escherichia coli; antigen; synthetic gene construction; mutein;
XX deletion mutation.
XX
OS Human immunodeficiency virus 1; group O isolate HAM112.
OS Synthetic.
XX
PN WO200004383-A2.
XX

```



```

PD 25-FEB-1999.
XX
XX 17-AUG-1998; 98WO-US017014.
XX
XX 15-AUG-1997; 97US-00911824.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;
XX WPI; 1999-190167/16.
XX DR N-PSDB; AAX56076.
XX
XX New isolated HIV-1 Group O env polypeptides - used for the detection of
XX PT anti-HIV antibodies and for the production of antibodies for use in
XX PT detection, purification and therapy.
XX
XX Claim 15; Fig 9; 138pp; English.
XX
XX The present invention describes (A) an isolated HIV-1 Group O env
XX CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env
XX CC polypeptide comprising an immunoreactive portion of a polypeptide as in
XX CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);
XX CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide
XX CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct
XX CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at
XX CC least one HIV-1 Group M env polypeptide; (5) an antigen construct
XX CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env
XX CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an
XX CC antigen construct comprising a first HIV-2 env polypeptide fused to a
XX CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as
XX CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a
XX CC host cell transformed by an expression vector as in (8); and (10) an
XX CC immunoassay kit for the detection of antibodies to HIV-1 comprising an
XX CC antigen construct as in (3)-(6). The antigen constructs can be used for
XX CC the detection of anti-HIV-1 antibodies in test samples. They can also be
XX CC used as immunogens to produce antibodies. The antibodies can be used to
XX CC purify HIV polypeptides, for therapy and for detection of HIV
XX CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 373 AA;
SQ
Query Match 92.0%; Score 162; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 1.2e-12;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRVNET 32
Db 117 ALETLLNQQLLDLWGCRGLVCYTSVRVNET 148
RESULT 13
AAY06979
ID AAY06979 standard; protein; 373 AA.
XX
XX AAY06979;
AC
XX
XX 06-JUL-1999 (first entry)
DT
XX
XX Recombinant pGO-11PL protein.
DE
XX
XX HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;
XX KW antibody; assay.
XX
XX Synthetic.
XX OS Human immunodeficiency virus 1.
XX
XX Key Location/Qualifiers
XX FH 2..46
XX FT Protein /note= "gpl20 sequence"
XX FT Peptide 47..245
XX FT /note= "gp41 sequence"
XX

```

---

```

PN WO9909410-A2.
XX
XX 25-FEB-1999.
XX
XX 07-AUG-1998; 98WO-US016506.
XX
XX 15-AUG-1997; 97US-00912129.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
XX PI Golden AM, Brennan CA, Devare SG;
XX DR WPI; 1999-190224/16.
XX PT N-PSDB; AAX37191.
XX
XX New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
XX PT be used in field assay, requiring no electricity and less specialised
XX PT equipment.
XX
XX Claim 1; Fig 9; 104pp; English.
XX
XX The invention relates to a rapid assay for simultaneous detection and
XX CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
XX CC method comprises (a) contacting the sample with a strip containing at
XX CC least one immobilised capture reagent per analyte and on which the sample
XX CC moves from the proximal to the distal end by capillary action, under
XX CC conditions sufficient to form capture reagent/analyte complexes, and (b)
XX CC determining the presence of analyte(s) by detecting a visible colour
XX CC change at the capture reagent site on the strip wherein the capture
XX CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
XX CC in AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
XX CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in
XX CC AAY06981. The invention is used to screen patients for antibodies to HIV-
XX CC 1 types O and M, and HIV-2. The invention will be particularly useful in
XX CC places and situation where equipment and/or electricity is not available.
XX CC The invention provides a screening method which is faster and requires
XX CC less equipment than prior art methods. The present sequence represents a
XX CC amino acid sequence of the recombinant pGO-11PL recombinant protein which
XX CC acts as a capture reagent for HIV-1 group O
XX
XX Sequence 373 AA;
SQ
Query Match 92.0%; Score 162; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 1.2e-12;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRVNET 32
Db 117 ALETLLNQQLLDLWGCRGLVCYTSVRVNET 148
RESULT 14
AAY09500
ID AAY09500 standard; protein; 460 AA.
XX
XX AAY09500;
AC
XX
XX 17-OCT-2003 (revised)
DT
XX 15-JUL-1999 (first entry)
DT
XX
XX HIV-1 Group O env polypeptide pGO-8CKs.
DE
XX
XX HIV; human immunodeficiency virus; antigen; detection; antibody;
XX KW differentiation; Group O; env; immunogen; immunoassay.
XX
XX Human immunodeficiency virus 1.
XX OS
XX
XX WO9909179-A2.
XX
XX 25-FEB-1999.
XX
XX 17-AUG-1998; 98WO-US017014.
XX

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 17.6134 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-13

Perfect score: 176

Sequence: 1 ALETLNQQLLDLWGRGLVCYTSVRWNET 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pap.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162	92.0	215	2	US-08-912-129A-58
2	162	92.0	245	2	US-08-912-129A-48
3	162	92.0	373	2	US-08-912-129A-52
4	162	92.0	460	2	US-08-912-129A-60
5	162	92.0	490	2	US-08-912-129A-50
6	162	92.0	618	2	US-08-912-129A-54
7	162	92.0	873	2	US-08-912-129A-61
8	159	90.3	116	4	US-09-462-917A-20
9	159	90.3	356	1	US-08-602-713-12
10	159	90.3	356	3	US-08-989-493-12
11	159	90.3	356	4	US-09-610-271-12
12	158	89.8	113	4	US-09-462-917A-2
13	158	89.8	117	4	US-09-462-917A-6
14	157	89.2	40	3	US-08-894-699-39
15	157	89.2	40	3	US-08-444-410-39
16	157	89.2	41	3	US-08-894-699-67
17	157	89.2	41	3	US-08-444-410-67
18	157	89.2	715	4	US-09-462-917A-134
19	156	88.6	37	4	US-08-817-441-86
20	156	88.6	37	4	US-08-817-441-94
21	156	88.6	104	4	US-08-817-441-100
22	156	88.6	113	4	US-09-462-917A-10
23	156	88.6	113	4	US-09-462-917A-12
24	156	88.6	200	3	US-08-965-056-104
25	156	88.6	862	4	US-09-206-551-15
26	155	88.1	40	3	US-08-894-699-37
27	155	88.1	40	3	US-08-894-699-42

28 155 88.1 40 3 US-09-444-410-37 Sequence 37, Appl  
29 155 88.1 40 3 US-09-444-410-42 Sequence 42, Appl  
30 155 88.1 110 4 US-09-462-917A-30 Sequence 30, Appl  
31 154 87.5 33 3 US-09-433-428D-10 Sequence 10, Appl  
32 154 87.5 40 3 US-08-894-699-40 Sequence 40, Appl  
33 154 87.5 40 3 US-09-444-410-40 Sequence 40, Appl  
34 154 87.5 113 4 US-09-462-917A-28 Sequence 28, Appl  
35 153 86.9 33 3 US-09-433-428D-4 Sequence 4, Appl  
36 153 86.9 33 3 US-09-433-428D-15 Sequence 15, Appl  
37 153 86.9 42 3 US-08-894-699-66 Sequence 66, Appl  
38 153 86.9 42 3 US-09-444-410-66 Sequence 66, Appl  
39 153 86.9 113 4 US-09-462-917A-18 Sequence 18, Appl  
40 153 86.9 116 4 US-09-462-917A-22 Sequence 22, Appl  
41 152 86.4 33 3 US-09-433-428D-5 Sequence 5, Appl  
42 152 86.4 33 3 US-09-433-428D-12 Sequence 12, Appl  
43 152 86.4 40 3 US-08-894-699-41 Sequence 41, Appl  
44 152 86.4 40 3 US-09-444-410-41 Sequence 41, Appl  
45 151 85.8 33 3 US-09-433-428D-1 Sequence 1, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-912-129A-58

; Sequence 58, Application US/08912129A

; Patent No. 5922533

; GENERAL INFORMATION:

; APPLICANT: VALLARI, ANADRUZELA S.

; APPLICANT: HACKETT, JOHN JR.

; APPLICANT: HICKMAN, ROBERT K.

; APPLICANT: VARITEK, VINCENT A. JR.

; APPLICANT: NECKLAWS, ELIZABETH A.

; APPLICANT: GOLDEN, ALAN M.

; APPLICANT: BRENNAN, CATHERINE A.

; APPLICANT: DEVARE, SUSHIL G.

; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO

; NUMBER OF SEQUENCES: 89

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: 100 Abbott Park Road

; CITY: Abbott Park

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS (Windows 95)

; SOFTWARE: Microsoft Word (ASCII format output)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/912,129A

; FILING DATE: 15-AUG-1997

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Danckers, Andreas M.

; REGISTRATION NUMBER: 32,652

; REFERENCE/DOCKET NUMBER: 6109 US 01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847-937-9803

; TELEFAX: 847-938-2623

; TELEX:

; INFORMATION FOR SEQ ID NO: 58:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 215 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-912-129A-58

```
Query Match          92.0%; Score 162; DB 2; Length 215;
Best Local Similarity 84.4%; Pred. No. 3.3e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLVCYTSVKWNET 148

RESULT 2
US-08-912-129A-48
; Sequence 48, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-129A-48

Query Match          92.0%; Score 162; DB 2; Length 245;
Best Local Similarity 84.4%; Pred. No. 3.8e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLVCYTSVKWNET 148

RESULT 3
US-08-912-129A-52
; Sequence 52, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
```

```
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-129A-52

Query Match          92.0%; Score 162; DB 2; Length 373;
Best Local Similarity 84.4%; Pred. No. 6.1e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLVCYTSVKWNET 148

RESULT 4
US-08-912-129A-60
; Sequence 60, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
```



STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckers, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-912-129A-60

Query Match 92.0%; Score 162; DB 2; Length 460;  
Best Local Similarity 84.4%; Pred. No. 7.7e-16;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||  
Db 362 ALETLLNQQLDLWGCRGLVCYTSVRWNET 393

RESULT 5  
US-08-912-129A-50  
Sequence 50, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997

CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckers, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 490 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-912-129A-50

Query Match 92.0%; Score 162; DB 2; Length 490;  
Best Local Similarity 84.4%; Pred. No. 8.3e-16;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||  
Db 362 ALETLLNQQLDLWGCRGLVCYTSVRWNET 393

RESULT 6  
US-08-912-129A-54  
Sequence 54, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckers, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 618 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-912-129A-54

Query Match 92.0%; Score 162; DB 2; Length 618;  
Best Local Similarity 84.4%; Pred. No. 1.1e-15;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32

Db 362 ALETLNQQLDLWGCRGLVCYTSVRWNET 393

## RESULT 7

US-08-912-129A-61  
; Sequence 61, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALIARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.  
; APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.  
; APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.  
; APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.  
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION

; NUMBER OF SEQUENCES: 89

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS (Windows 95)

; SOFTWARE: Microsoft Word (ASCII format output)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/912,129A

; FILING DATE: 15-AUG-1997

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Dankers, Andreas M.

; REGISTRATION NUMBER: 32,652

; REFERENCE/DOCKET NUMBER: 6109 US.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847-937-9803

; TELEFAX: 847-938-2623

; TELEX:

; INFORMATION FOR SEQ ID NO: 61:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 873 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Protein

US-08-912-129A-61

Query Match 92.0%; Score 162; DB 2; Length 873;  
Best Local Similarity 84.4%; Pred. No. 1.1e-15;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32

Db 591 ALETLNQQLDLWGCRGLVCYTSVRWNET 622

## RESULT 8

US-09-462-917A-20

; Sequence 20, Application US/09462917A

; Patent No. 6511801

; GENERAL INFORMATION:

; APPLICANT: Delaporte, Eric

; APPLICANT: Peeters, Martine

; APPLICANT: Saman, Eric

; APPLICANT: Vanden Haesevelde, Marlen

; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof

; FILE REFERENCE: INNS:014 11362.0014.NFUS00

; CURRENT APPLICATION NUMBER: US/09/462,917A

; PRIOR FILING DATE: 2000-04-03

; PRIOR APPLICATION NUMBER: PCT/EP98/04522

; NUMBER OF SEQ ID NOS: 152

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 20

; LENGTH: 116

; TYPE: PRT

; ORGANISM: Human

US-09-462-917A-20

Query Match 90.3%; Score 159; DB 4; Length 116;

Best Local Similarity 84.4%; Pred. No. 4.6e-16;

Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32

Db 29 ALETLNQQLDLWGCRGLVCYTSVRWNET 60

## RESULT 9

US-08-602-713-12

; Sequence 12, Application US/08602713

; Patent No. 5798205

; GENERAL INFORMATION:

; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;

; APPLICANT: G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;

; APPLICANT: Zekeng, L opold Achengui

; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use

; TITLE OF INVENTION: (MVP-2901/94)

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/602,713

; FILING DATE: 16-FEBRUARY-1996

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 195 05 262

; FILING DATE: 16-FEBRUARY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5798205man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LEDEB 203

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
US-08-602-713-12

Query Match 90.3%; Score 159; DB 1; Length 356;  
Best Local Similarity 81.2%; Pred. No. 1.6e-15;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 312 ALETLNQQLDLWGCRGLVCYTSVRWNET 343

RESULT 10  
US-08-989-493-12  
; Sequence 12, Application US/08989493  
; Patent No. 6162631  
; GENERAL INFORMATION:  
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;  
; APPLICANT: G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;  
; APPLICANT: Zekeng, L opold Achengui  
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989,493  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/602,713  
FILING DATE: 16-FEBRUARY-1996  
APPLICATION NUMBER: DE 195 05 262  
FILING DATE: 16-FEBRUARY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6162631man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LEDER 203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
US-08-989-493-12

Query Match 90.3%; Score 159; DB 3; Length 356;  
Best Local Similarity 81.2%; Pred. No. 1.6e-15;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32

Db 312 ALETLNQQLDLWGCRGLVCYTSVRWNET 343  
|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11  
US-09-610-271-12  
; Sequence 12, Application US/09610271  
; Patent No. 6548635  
; GENERAL INFORMATION:  
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;  
; APPLICANT: G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;  
; APPLICANT: Zekeng, L opold Achengui  
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,271  
FILING DATE: 06-Jul-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/602,713  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6548635man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LEDER 203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-610-271-12

Query Match 90.3%; Score 159; DB 4; Length 356;  
Best Local Similarity 81.2%; Pred. No. 1.6e-15;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 312 ALETLNQQLDLWGCRGLVCYTSVRWNET 343

RESULT 12  
US-09-462-917A-2  
; Sequence 2, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NFUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A





THE PUBLIC WORKS (CIVIL)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 44.7731 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-13  
Perfect score: 176  
Sequence: 1 ALETLLNQQLLDWGCGRGLVCYTSVRWNET 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/FCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162	92.0	215	8	US-08-911-824-58
2	162	92.0	245	8	US-08-911-824-48
3	162	92.0	281	8	US-08-911-824-120
4	162	92.0	373	8	US-08-911-824-52
5	162	92.0	460	8	US-08-911-824-60
6	162	92.0	488	8	US-08-911-824-95
7	162	92.0	490	8	US-08-911-824-50
8	162	92.0	526	8	US-08-911-824-97
9	162	92.0	618	8	US-08-911-824-54
10	162	92.0	706	8	US-08-911-824-93
11	162	92.0	736	8	US-08-911-824-91
12	162	92.0	873	8	US-08-911-824-61
13	159	90.3	116	14	US-10-320-786-20
14	159	90.3	356	14	US-10-357-400-12
15	158	89.8	113	14	US-10-320-786-2

16	158	89.8	117	14	US-10-320-786-6	Sequence 6, Appli
17	157	89.2	715	14	US-10-320-786-134	Sequence 134, App
18	156	88.6	37	14	US-10-026-741-86	Sequence 86, Appl
19	156	88.6	37	14	US-10-026-741-94	Sequence 94, Appl
20	156	88.6	104	14	US-10-026-741-100	Sequence 100, Appl
21	156	88.6	113	14	US-10-320-786-10	Sequence 10, Appl
22	156	88.6	113	14	US-10-320-786-12	Sequence 12, Appl
23	156	88.6	200	9	US-09-854-816-104	Sequence 104, App
24	156	88.6	862	14	US-10-369-294-15	Sequence 15, Appl
25	155	88.1	110	14	US-10-320-786-30	Sequence 30, Appl
26	154	87.5	113	14	US-10-320-786-28	Sequence 28, Appl
27	153	86.9	113	14	US-10-320-786-18	Sequence 18, Appl
28	153	86.9	116	14	US-10-320-786-22	Sequence 22, Appl
29	151	85.8	113	14	US-10-320-786-24	Sequence 24, Appl
30	149	84.7	113	14	US-10-320-786-40	Sequence 40, Appl
31	148	84.1	35	14	US-10-026-741-101	Sequence 101, App
32	148	84.1	37	14	US-10-026-741-7	Sequence 7, Appli
33	148	84.1	37	14	US-10-026-741-88	Sequence 88, Appl
34	148	84.1	37	14	US-10-026-741-90	Sequence 90, Appl
35	148	84.1	110	14	US-10-320-786-36	Sequence 36, Appl
36	148	84.1	351	14	US-10-026-741-47	Sequence 47, Appl
37	148	84.1	877	14	US-10-026-741-102	Sequence 102, App
38	145	82.4	32	14	US-10-364-360-30	Sequence 30, Appl
39	145	82.4	110	14	US-10-320-786-14	Sequence 14, Appl
40	144	81.8	110	14	US-10-320-786-16	Sequence 16, Appl
41	144	81.8	110	14	US-10-320-786-38	Sequence 38, Appl
42	143	81.2	35	9	US-09-886-156-62	Sequence 62, Appl
43	143	81.2	35	9	US-09-886-150-62	Sequence 62, Appl
44	143	81.2	35	10	US-09-886-149-62	Sequence 62, Appl
45	143	81.2	35	10	US-09-886-159-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911.824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 92.0%; Score 162; DB 8; Length 215;  
Best Local Similarity 84.4%; Pred. No. 2.7e-14;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALETLLNQQLLDWGCGRGLVCYTSVRWNET 32  
Db 117 ALETLLNQQLLDWGCGRGLVCYTSVRWNET 148

RESULT 2

US-08-911-824-48

```
; Sequence 48, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL
US-08-911-824-48

Query Match          92.0%; Score 162; DB 8; Length 245;
Best Local Similarity 84.4%; Pred. No. 3.1e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLICYTSVKWNET 148

RESULT 3
US-08-911-824-120
; Sequence 120, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 120
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL
US-08-911-824-120

Query Match          92.0%; Score 162; DB 8; Length 281;
Best Local Similarity 84.4%; Pred. No. 3.5e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLICYTSVKWNET 148

; Sequence 48, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL
US-08-911-824-52

Query Match          92.0%; Score 162; DB 8; Length 373;
Best Local Similarity 84.4%; Pred. No. 4.7e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 ALETLNQQLLDLWGCGRGLICYTSVKWNET 148

RESULT 5
US-08-911-824-60
; Sequence 60, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS
US-08-911-824-60

Query Match          92.0%; Score 162; DB 8; Length 460;
Best Local Similarity 84.4%; Pred. No. 5.8e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
    |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 362 ALETLNQQLLDLWGCGRGLICYTSVKWNET 393

RESULT 6
US-08-911-824-95
; Sequence 95, Application US/08911824
; Publication No. US20030004323A1
```



; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories  
 ; APPLICANT: Hackett, John R., Jr.  
 ; APPLICANT: Yamaguchi, Julie  
 ; APPLICANT: Golden, Alan M.  
 ; APPLICANT: Brennan, Catherine A.  
 ; APPLICANT: Hickman, Robert K.  
 ; APPLICANT: Devare, Sushil G.  
 ; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
 ; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
 ; FILE REFERENCE: 6165 US 01  
 ; CURRENT APPLICATION NUMBER: US/08/911,824  
 ; CURRENT FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 121  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 95  
 ; LENGTH: 488  
 ; TYPE: PRT  
 ; ORGANISM: Human Immunodeficiency Virus  
 ; FEATURE:  
 ; OTHER INFORMATION: Encodes recombinant protein pGO-14PL  
 US-08-911-824-95

Query Match 92.0%; Score 162; DB 8; Length 488;  
 Best Local Similarity 84.4%; Pred. No. 6.2e-14;  
 Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
 |||||:|||||:|||||:|||||:|||||:|||||  
 Db 117 ALETLNQQLDLWGCRGLVCYTSVRWNET 148

RESULT 7  
 US-08-911-824-50  
 ; Sequence 50, Application US/08911824  
 ; Publication No. US20030004323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories  
 ; APPLICANT: Hackett, John R., Jr.  
 ; APPLICANT: Yamaguchi, Julie  
 ; APPLICANT: Golden, Alan M.  
 ; APPLICANT: Brennan, Catherine A.  
 ; APPLICANT: Hickman, Robert K.  
 ; APPLICANT: Devare, Sushil G.  
 ; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
 ; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
 ; FILE REFERENCE: 6165 US 01  
 ; CURRENT APPLICATION NUMBER: US/08/911,824  
 ; CURRENT FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 121  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 50  
 ; LENGTH: 490  
 ; TYPE: PRT  
 ; ORGANISM: Human Immunodeficiency Virus  
 ; FEATURE:  
 ; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS  
 US-08-911-824-50

Query Match 92.0%; Score 162; DB 8; Length 490;  
 Best Local Similarity 84.4%; Pred. No. 6.2e-14;  
 Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
 |||||:|||||:|||||:|||||:|||||:|||||  
 Db 362 ALETLNQQLDLWGCRGLVCYTSVRWNET 393

RESULT 8  
 US-08-911-824-97  
 ; Sequence 97, Application US/08911824  
 ; Publication No. US20030004323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories

; APPLICANT: Abbott Laboratories  
 ; APPLICANT: Hackett, John R., Jr.  
 ; APPLICANT: Yamaguchi, Julie  
 ; APPLICANT: Golden, Alan M.  
 ; APPLICANT: Brennan, Catherine A.  
 ; APPLICANT: Hickman, Robert K.  
 ; APPLICANT: Devare, Sushil G.  
 ; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
 ; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
 ; FILE REFERENCE: 6165 US 01  
 ; CURRENT APPLICATION NUMBER: US/08/911,824  
 ; CURRENT FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 121  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 97  
 ; LENGTH: 526  
 ; TYPE: PRT  
 ; ORGANISM: Human Immunodeficiency Virus  
 ; FEATURE:  
 ; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS  
 US-08-911-824-97

Query Match 92.0%; Score 162; DB 8; Length 526;  
 Best Local Similarity 84.4%; Pred. No. 6.7e-14;  
 Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
 |||||:|||||:|||||:|||||:|||||:|||||  
 Db 362 ALETLNQQLDLWGCRGLVCYTSVRWNET 393

RESULT 9  
 US-08-911-824-54  
 ; Sequence 54, Application US/08911824  
 ; Publication No. US20030004323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories  
 ; APPLICANT: Hackett, John R., Jr.  
 ; APPLICANT: Yamaguchi, Julie  
 ; APPLICANT: Golden, Alan M.  
 ; APPLICANT: Brennan, Catherine A.  
 ; APPLICANT: Hickman, Robert K.  
 ; APPLICANT: Devare, Sushil G.  
 ; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
 ; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
 ; FILE REFERENCE: 6165 US 01  
 ; CURRENT APPLICATION NUMBER: US/08/911,824  
 ; CURRENT FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 121  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 54  
 ; LENGTH: 618  
 ; TYPE: PRT  
 ; ORGANISM: Human Immunodeficiency Virus  
 ; FEATURE:  
 ; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS  
 US-08-911-824-54

Query Match 92.0%; Score 162; DB 8; Length 618;  
 Best Local Similarity 84.4%; Pred. No. 7.9e-14;  
 Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32  
 |||||:|||||:|||||:|||||:|||||:|||||  
 Db 362 ALETLNQQLDLWGCRGLVCYTSVRWNET 393

RESULT 10  
 US-08-911-824-93  
 ; Sequence 93, Application US/08911824  
 ; Publication No. US20030004323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories

```
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-13CKs
US-08-911-824-93

Query Match          92.0%; Score 162; DB 8; Length 706;
Best Local Similarity 84.4%; Pred. No. 9e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 608 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 639

RESULT 11
US-08-911-824-91
; Sequence 91, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-12CKs
US-08-911-824-91

Query Match          92.0%; Score 162; DB 8; Length 736;
Best Local Similarity 84.4%; Pred. No. 9.4e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 608 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 639

RESULT 12
US-08-911-824-61
; Sequence 61, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
```

```
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.01
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: HIV-1 Group O isolate HAM112
US-08-911-824-61

Query Match          92.0%; Score 162; DB 8; Length 873;
Best Local Similarity 84.4%; Pred. No. 1.1e-13;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 591 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 622

RESULT 13
US-10-320-786-20
; Sequence 20, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; OTHER INFORMATION:
US-10-320-786-20

Query Match          90.3%; Score 159; DB 14; Length 116;
Best Local Similarity 84.4%; Pred. No. 3.8e-14;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 32
|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 ALETLNQQLLDLWGCGRGLVCYTSVRWNET 60

RESULT 14
US-10-357-400-12
; Sequence 12, Application US/10357400
; Publication No. US20030147917A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
; G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;
; Zekeng, L opold Achengui
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use
; (MVP-2901/94)
```

NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/357,400  
FILING DATE: 04-Feb-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,271  
FILING DATE: 06-Jul-2000  
APPLICATION NUMBER: 08/602,713  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. US20030147917Alman D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LEDER 203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-357-400-12

Query Match 90.3%; Score 159; DB 14; Length 356;  
Best Local Similarity 81.2%; Pred. No. 1.2e-13;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 312 ALETLMNQQLLDLWGCKGKLCYTSVKWNET 343

RESULT 15  
US-10-320-786-2  
Sequence 2, Application US/10320786  
Publication No. US20030180759A1  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
APPLICANT: Vanden Haesevelde, Marlen  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: US 09/462,917  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 113  
TYPE: PRT  
ORGANISM: Human  
US-10-320-786-2

Query Match 89.8%; Score 158; DB 14; Length 113;  
Best Local Similarity 81.2%; Pred. No. 5e-14;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ALETLLNQQLLDLWGCRGLVCYTSVRWNET 32  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 29 ALETLIQNQQLLDLWGCKGRIVCYTSVKWNDT 60

Search completed: May 7, 2004, 18:29:22  
Job time : 45.7731 secs

THIS PAGE BLANK (2510)

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	156	88.6	104	2	S52930	GP41 ENV protein -	
2	156	88.6	853	2	A53034	gag polyprotein -	
3	148	84.1	877	2	S49197	envelope protein p	
4	115	65.3	854	1	VCLJUS1	env polyprotein pr	
5	108	61.4	357	2	S21990	envelope protein g	
6	108	61.4	358	2	S22002	envelope protein g	
7	108	61.4	454	2	B41621	env polyprotein D	
8	107	60.8	358	2	S22000	envelope protein g	
9	107	60.8	358	2	S70417	envelope protein g	
10	106	60.2	443	2	C41621	env polyprotein p	
11	106	60.2	853	2	S54384	envelope polyprote	
12	106	60.2	855	1	VCLJZR	env polyprotein pr	
13	105	59.7	357	2	S22006	envelope protein g	
14	105	59.7	357	2	S21994	envelope protein g	
15	105	59.7	357	2	S22004	envelope protein g	
16	105	59.7	357	2	S21996	envelope protein g	
17	105	59.7	357	2	S21992	envelope protein g	
18	105	59.7	358	2	S21998	envelope protein g	
19	105	59.7	445	2	A41621	env polyprotein M	
20	105	59.7	843	1	H44001	env polyprotein pr	
21	105	59.7	852	1	T120JB	env polyprotein -	
22	105	59.7	852	2	T12016	envelope glycoprot	
23	105	59.7	854	2	S13288	env protein - huma	
24	105	59.7	855	1	VCLJA2	env polyprotein pr	
25	105	59.7	856	1	VCLJH3	env polyprotein pr	
26	105	59.7	856	1	VCLJVL	env polyprotein pr	
27	105	59.7	856	1	VCLJ3W	env polyprotein pr	
28	105	59.7	861	1	VCLJLV	env polyprotein pr	
29	105	59.7	861	1	VCLJSC	env polyprotein pr	



[illegible]

RESULT 12  
VCLJZR  
env polyprotein precursor - human immunodeficiency virus Zr-6  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus Zr-6  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
C:Accession: D26192  
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; C  
Gene 52, 71-82, 1987  
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti  
A:Reference number: A26192; MUID:87248097; PMID:3036660  
A:Accession: D26192  
A:Molecule type: DNA  
A:Residues: 1-855 <SRI>  
A:Cross-references: GB:K03458; GB:M16322; NID:G329398; PIDN:AAA45380.1; PID:G329403  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-855/Product: env polyprotein #status predicted <MAT>  
F:20-500/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:501-855/Product: transmembrane glycoprotein #status predicted <TM>  
F:87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,404

Query Match 60.2%; Score 106; DB 1; Length 855;  
Best Local Similarity 53.1%; Pred. No. 7.6e-07;  
Matches 17; Conservative 8; Mismatches 7; Indels 0; Gaps 0;  
QY 1 ALETLNQOQLDLWGCGRLVCYTSVRNWT 32  
DB 581 AVERYLKQQLLGWCGSGKLICTTTVPWNSS 612

RESULT 13  
S22006  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
R:Accession: S70420; S22006  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70420  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <ST2>  
A:Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191  
A:Experimental source: patient L  
A:Note: submitted to the EMBL Data Library, July 1991  
C:Superfamily: type E retrovirus env polyprotein

Query Match 59.7%; Score 105; DB 2; Length 357;  
Best Local Similarity 56.7%; Pred. No. 4.5e-07;  
Matches 17; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ALETLNQOQLDLWGCGRLVCYTSVRWN 30  
DB 83 AVERYLKQQLLGWCGSGKLICTTTAVPNW 112

RESULT 14  
S21994  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 27B  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21994; S70421  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin

A:Reference number: S21990  
A:Accession: S21994  
A:Molecule type: DNA  
A:Residues: 1-357 <ST2>  
A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70421  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-140,'X',142-312,'X',314-357 <ST2>  
A:Cross-references: EMBL:X61355; NID:g60179  
C:Superfamily: type E retrovirus env polyprotein

Query Match 59.7%; Score 105; DB 2; Length 357;  
Best Local Similarity 56.7%; Pred. No. 4.5e-07;  
Matches 17; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ALETLNQOQLDLWGCGRLVCYTSVRWN 30  
DB 83 AVERYLKQQLLGWCGSGKLICTTTVPWN 112

RESULT 15  
S22004  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 4B  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22004; S70419  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determin  
A:Reference number: S21990  
A:Accession: S22004  
A:Molecule type: DNA  
A:Residues: 1-357 <ST2>  
A:Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70419  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-292,'X',294-357 <ST2>  
A:Cross-references: EMBL:X61353; NID:g60188  
C:Superfamily: type E retrovirus env polyprotein

Query Match 59.7%; Score 105; DB 2; Length 357;  
Best Local Similarity 56.7%; Pred. No. 4.5e-07;  
Matches 17; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ALETLNQOQLDLWGCGRLVCYTSVRWN 30  
DB 83 AVERYLKQQLLGWCGSGKLICTTTVPWN 112

Search completed: May 7, 2004, 17:51:03  
Job time : 13.4454 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 8.06723 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-13  
Perfect score: 176  
Sequence: 1 ALETLQOQLLDLWGCGRGLVCYTSVRWNET 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	65.3	854	1 ENV_SIVCZ	P17281 chimpanzee
2	106	60.2	853	1 ENV_HV1Z2	P12487 human immun
3	106	60.2	855	1 ENV_HV1Z6	P04580 human immun
4	105	59.7	843	1 ENV_HV1Y2	P35961 human immun
5	105	59.7	847	1 ENV_HV1S1	P19550 human immun
6	105	59.7	847	1 ENV_HV1W2	P05880 human immun
7	105	59.7	851	1 ENV_HV1B8	P04582 human immun
8	105	59.7	852	1 ENV_HV1B9	P12488 human immun
9	105	59.7	852	1 ENV_HV1S3	P19549 human immun
10	105	59.7	853	1 ENV_HV1MF	P19551 human immun
11	105	59.7	855	1 ENV_HV1A2	P03378 human immun
12	105	59.7	855	1 ENV_HV1OY	P20888 human immun
13	105	59.7	855	1 ENV_HV1B1	P03375 human immun
14	105	59.7	856	1 ENV_HV1H2	P04578 human immun
15	105	59.7	856	1 ENV_HV1LW	Q70626 human immun
16	105	59.7	856	1 ENV_HV1PV	P03376 human immun
17	105	59.7	856	1 ENV_HV1SC	P05878 human immun
18	105	59.7	856	1 ENV_HV1W1	P31872 human immun
19	105	59.7	861	1 ENV_HV1BR	P03377 human immun
20	105	59.7	865	1 ENV_HV1J3	P04579 human immun
21	105	59.7	867	1 ENV_HV1J3	P12489 human immun
22	103	58.5	848	1 ENV_HV1JR	P20871 human immun
23	103	58.5	856	1 ENV_HV1H3	P04624 human immun
24	103	58.5	856	1 ENV_HV1MN	P05877 human immun
25	103	58.5	856	1 ENV_HV1ZH	P05881 human immun
26	103	58.5	868	1 ENV_HV1C4	P05879 human immun
27	102	58.0	846	1 ENV_HV1ND	P18799 human immun
28	101	57.4	863	1 ENV_HV1Z8	P05882 human immun
29	99	56.2	853	1 ENV_HV1EL	P04581 human immun
30	99	56.2	861	1 ENV_HV1KB	P31819 human immun
31	96	54.5	859	1 ENV_HV1MA	P04583 human immun
32	87	49.4	821	1 ENV_SIVGB	P22380 simian immu
33	84.5	48.0	856	1 ENV_HV2N2	P05883 human immun

RESULT 1				
ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retrovirdae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."			
RL	Nature 345:356-359(1990).			
CC	-1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; X52154; CAA36407.1; -			
DR	PIR; S09990; VCLJJSI.			
DR	HIV; X52154; ENVSCPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. .) (POTENTIAL).

ALIGNMENTS

```

FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 65.3%; Score 115; DB 1; Length 854;
Best Local Similarity 59.4%; Pred. No. 7, 8e-10;
Matches 19; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALETLLNQQLDLWGCGRLVCYTSVRWNET 32
Db 572 AVERYLDQQLLGWCGSKAVCYTTPWNSS 603

RESULT 2
ENV_HV122
ID ENV_HV122 STANDARD; PRT; 853 AA.
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Z22/CDC-Z34 isolate) (HIV-1).
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV3226.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```

FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B0B8CBAFF7008 CRC64;

Query Match 60.2%; Score 106; DB 1; Length 853;
Best Local Similarity 53.1%; Pred. No. 2e-08;
Matches 17; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALETLLNQQLDLWGCGRLVCYTSVRWNET 32
Db 579 AVERYLDQQLLGWCGSKLICITTPWNSS 610

RESULT 3
ENV_HV126
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RA MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.

```



FT	CARBOHYD	382	382	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	388	388	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	398	398	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	401	401	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	438	438	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	454	454	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	602	602	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	607	607	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	628	628	N-LINKED (GLCNAC..)	(POTENTIAL).
SQ	SEQUENCE	847 AA;	96135 MW;	OA901317DF7FFZAB	CRG64;

Query Match            59.7%; Score 105; DB 1; Length 847;  
Best Local Similarity   56.7%; Pred. No. 2.8e-08;  
Matches 17; Conservative

Qy	1	ALETLLNQQLDLWGCRGLVCYTVSRWN	30
	:	: : : : : : : : : : : : :	:
Db	573	AVERYLKDQLLGIWGCGLICTTAVPMN	602

```

RESULT 6
ENV_HV1W2      STANDARD;          PRT;    847 AA.
AC   P05880;
AT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DD   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DN   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
OS   GN.
OC   Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OX   Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11705;
[1]
RN   SEQUENCE FROM N.A.
RX   MEDLINE=86235450; PubMed=3012778;
RA   Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA   Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.; or
RT   "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT   at risk for AIDS."
RL   Science 232:1548-1553(1986).
CC   -I- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC   BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC   WAS PERINATALLY INFECTED BY HER MOTHER.
-----
CC   THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to licens@isb-sib.ch).
-----
CC   EMBL; M12507; AAB12990.1; -.
DR   HIV; M12507; ENV$WMJ2.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL              1
FT   CHAIN               30 501
FT   FT CHAIN           502 847
FT   DISULFID            53 73
FT   DISULFID           118 202
FT   DISULFID           125 193
FT   DISULFID           130 152
FT   DISULFID           215 244
FT   DISULFID           225 236
FT   DISULFID           293 326
EX   EXTERIOR MEMBRANE GLYCOPROTEIN.
EX   TRANSMEMBRANE GLYCOPROTEIN.
BY   SIMILARITY.
BY   SIMILARITY.
BY   SIMILARITY.
BY   SIMILARITY.
BY   SIMILARITY.
BY   SIMILARITY.
```

[illegible]



FT	CHAIN		31	509	EXTERIOR MEMBRANE GLYCOPROTEIN.				
FT	CHAIN		510	853	TRANSMEMBRANE GLYCOPROTEIN.				
FT	DISULFID		54	74	BY SIMILARITY.				
FT	DISULFID		119	203	BY SIMILARITY.				
FT	DISULFID		126	194	BY SIMILARITY.				
FT	DISULFID		131	157	BY SIMILARITY.				
FT	DISULFID		216	245	BY SIMILARITY.				
FT	DISULFID		294	329	BY SIMILARITY.				
FT	DISULFID		376	443	BY SIMILARITY.				
FT	DISULFID		383	416	BY SIMILARITY.				
FT	CARBOHYD		88	88	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		136	136	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		141	141	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		156	156	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		160	160	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		186	186	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		195	195	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		232	232	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		239	239	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		330	330	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		334	354	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		384	384	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		390	390	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		395	395	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		404	404	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		446	446	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		461	461	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		609	609	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		614	614	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		623	623	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		635	635	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		672	672	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		748	748	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD		814	814	N-LINKED (GLCNAC. .) (POTENTIAL).				
SQ	SEQUENCE		853 AA;	96912 MW; 3377B933B6F722ABA CRC64;					
					Query Match 59.7%; Score 105; DB 1; Length 853;				
					Best Local Similarity 56.7%; Pred. No. 2.8e-08;				
					Matches 17; Conservative 7; Mismatches 6; Indels 0;				
OY	1 ALETLLQNQLLDLWGCGRGLVCTSVRN 30				:   ::     :     :   :   :   :				
Db	580 AVERYLKDQLLGWGCSGKLICTTAPWN 609				:   ::     :     :   :   :   :				
<hr/>									
RESULT 11									
ID	ENV_HVIA2	STANDARD;	PRT;	855 AA.					
AC	P03378;								
DT	21-JUL-1986 (Rel. 01, Created)								
DT	21-JUL-1986 (Rel. 01, Last sequence update)								
DT	15-JUL-1999 (Rel. 36, Last annotation update)								
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membran glycoprotein (GP120); Transmembrane glycoprotein (GP41)].								
DE	ENV.								
OS	Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).								
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.								
OX	NCBI_TaxId=11685;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=85090453; PubMed=2578227;								
RA	Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,								
RA	Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph J.								
RA	Levy J.A., Dina D., Luciw P.A.;								
RT	"Nucleotide sequence and expression of an AIDS-associated retroviral (ARV-2).";								
RL	Science 227:484-492(1985).								



-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; K02007; AAB59882.1; -;  
PIR; A03976; VCLJAJ2.  
HIV; K02007; ENV5SF2.  
InterPro; IPR000328; Env GP41.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
KW SIGNAL.  
FT CHAIN 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 208 BY SIMILARITY.  
FT DISULFID 125 199 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 221 250 BY SIMILARITY.  
FT DISULFID 231 242 BY SIMILARITY.  
FT DISULFID 299 333 BY SIMILARITY.  
FT DISULFID 380 442 BY SIMILARITY.  
FT DISULFID 387 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 59.7%; Score 105; DB 1; Length 855;  
Best Local Similarity 56.7%; Pred. No. 2.8e-08;  
Matches 17; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALETLNQQLDLWGCRGLVCYTSVRWN 30  
Db 581 AVERLYRDOQLLGWGCSGLKICTTAVPNW 610

RESULT 12

ENV\_HV10Y  
ID ENV\_HV10Y STANDARD; PRT; 855 AA.  
AC P20888;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor (Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
GN ENV.  
OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90148544; PubMed=2559749;  
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;  
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese  
individual presenting an atypical western blot.";  
RL AIDS 3:707-715 (1989).  
CC -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A  
HEALTHY GABONESE INDIVIDUAL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M26727; AAA83397.1; -;  
HIV; M26727; ENV5OYI.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
KW SIGNAL.  
FT CHAIN 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 210 BY SIMILARITY.  
FT DISULFID 125 201 BY SIMILARITY.  
FT DISULFID 130 162 BY SIMILARITY.  
FT DISULFID 223 252 BY SIMILARITY.  
FT DISULFID 233 244 BY SIMILARITY.  
FT DISULFID 301 335 BY SIMILARITY.  
FT DISULFID 381 442 BY SIMILARITY.  
FT DISULFID 388 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).







FT	CARBOHYD	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA; 96938 MW; 0C241332CF7E6687 CRC64;		

Query Match 59.7%; Score 105; DB 1; Length 856;

Best Local Similarity 56.7%; Pred. No. 2.8e-08;

Matches 17; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ALETLQOQLDLWGCGRGLVCYTSVRWN 30

Db 582 AVERYLKDOQLGIWGCGRGLICTTAVPNW 611

Search completed: May 7, 2004, 17:43:57

Job time : 9.06723 secs

COLORED MAN / 14 EIGHT SIX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 40.605 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-13

Perfect score: 176  
Sequence: 1 ALETLQNQQLDLWGCRCRLVCYTSVRWNET 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162	92.0	124	15	Q9IHU7 human immun
2	162	92.0	126	15	Q9IHV1 human immun
3	162	92.0	172	15	Q9IEB3 human immun
4	162	92.0	234	15	Q9IEC2 human immun
5	161	91.5	209	15	Q9IEE6 human immun
6	161	91.5	240	15	Q9IE32 human immun
7	160	90.9	125	15	Q9IHU8 human immun
8	160	90.9	216	15	Q9IEA5 human immun
9	160	90.9	230	15	Q9IEB2 human immun
10	160	90.9	536	15	Q9IEE5 human immun
11	160	90.9	544	15	Q9IED9 human immun
12	159	90.3	116	15	O40459 human immun
13	159	90.3	137	15	Q9IHV5 human immun
14	159	90.3	155	15	Q8J3Q4 human immun
15	159	90.3	208	15	Q9IEA3 human immun
16	159	90.3	238	15	Q9DIK1 human immun

17	159	90.3	342	15	O11942	human immun
18	159	90.3	418	15	O36547	human immun
19	159	90.3	871	15	O57074	human immun
20	159	90.3	871	15	O8Q7I2	human immun
21	159	90.3	876	15	O8Q7H3	human immun
22	158	89.8	105	15	O11939	human immun
23	158	89.8	114	15	O40472	human immun
24	158	89.8	118	15	O40451	human immun
25	158	89.8	213	15	Q9IEC4	human immun
26	158	89.8	235	15	Q9IE54	human immun
27	158	89.8	532	15	Q9IEF0	human immun
28	158	89.8	872	15	Q900Y5	human immun
29	158	89.8	880	15	Q8Q7H1	human immun
30	157	89.2	116	15	O40458	human immun
31	157	89.2	134	15	Q9IHV4	human immun
32	157	89.2	158	15	Q8J3N6	human immun
33	157	89.2	183	15	Q9IEC1	human immun
34	157	89.2	216	15	Q9IEC5	human immun
35	157	89.2	220	15	Q9IEC7	human immun
36	157	89.2	225	15	Q9IEA0	human immun
37	157	89.2	512	15	Q9IED2	human immun
38	157	89.2	535	15	Q9IEF2	human immun
39	157	89.2	545	15	Q9IED7	human immun
40	157	89.2	875	15	Q8Q7H4	human immun
41	157	89.2	879	15	Q9WIU9	human immun
42	157	89.2	900	15	Q9QN28	human immun
43	156	88.6	104	15	O76163	human immun
44	156	88.6	114	15	O40457	human immun
45	156	88.6	114	15	O40457	human immun

ALIGNMENTS

RESULT 1

Q9IHU7 PRELIMINARY; PRT; 124 AA.  
AC Q9IHU7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97CM768;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,  
RA Pieniazek D., Schable C., Lal R.B.;  
RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O.";  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
DR EMBL; AF229237; AAF71914.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
FT Transmembrane.  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 15187 MW; 105D515F114450F8 CRC64;

Query Match: 92.0%; Score 162; DB 15; Length 124;  
Best Local Similarity 84.4%; Pred. NO. 1.3e-16;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLQNQQLDLWGCRCRLVCYTSVRWNET 32

DB 23 ALETLQNQQLDLWGCRCRLVCYTSVRWNET 54

```
RESULT 2
Q9IHV1 ID Q9IHV1 PRELIMINARY; PRT; 126 AA.
AC Q9IHV1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=97ES203;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF292333; AAF71910.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 126
FT NON_TER 126
SQ SEQUENCE 126 AA; 15169 MW; 13FB101ECDCF0DDD CRC64;

Query Match 92.0%; Score 162; DB 15; Length 126;
Best Local Similarity 87.5%; Pred. No. 1.4e-16;
Matches 28; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32
DB 30 ALETLNQQLDLWGCRGLVCYTSVRWNETS 61

RESULT 3
Q9IEB3 ID Q9IEB3 PRELIMINARY; PRT; 172 AA.
AC Q9IEB3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=BCF99;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236407; CAB96255.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 172
FT NON_TER 172
SQ SEQUENCE 172 AA; 20388 MW; 972C7AF2A7546B13 CRC64;
```

```
Query Match 92.0%; Score 162; DB 15; Length 172;
Best Local Similarity 84.4%; Pred. No. 1.9e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32
DB 26 ALETLNQQLDLWGCRGLVCYTSVRWNET 57

RESULT 4
Q9IEC2 ID Q9IEC2 PRELIMINARY; PRT; 234 AA.
AC Q9IEC2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=BCF100;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236397; CAB96246.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 234
FT NON_TER 234
SQ SEQUENCE 234 AA; 27036 MW; 39B050B3F8555A8C CRC64;

Query Match 92.0%; Score 162; DB 15; Length 234;
Best Local Similarity 84.4%; Pred. No. 2.6e-16;
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLNQQLDLWGCRGLVCYTSVRWNET 32
DB 48 ALETLNQQLDLWGCRGLVCYTSVRWNET 79

RESULT 5
Q9IE66 ID Q9IE66 PRELIMINARY; PRT; 209 AA.
AC Q9IE66;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide, gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=BCF107;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauchere P.,
RA Depienne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ238862; CAB96300.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
```

KW	Transmembrane..					
FT	NON_TER	1				
FT	NON_TER	209				
SQ	SEQUENCE	209 AA;	24116 MW;	91AC9BE2BFE74B83 CRC64;		
 Query Match						
Best Local Similarity 91.5%; Score 161; DB 15; Length 209;						
Matches 27; Conservative 4; Mismatches 1; Indels 0; Gaps 0;						
OY	1 ALETLNQQLLDLWGCRGLVCYTYSVRWNKT 32					
DB	:     :     :     :     :					
DB	46 ALETLIQNQQLLSLWGCKGRILCYTSVKWNKT 77					
 RESULT 6						
ID	Q9IE32	PRELIMINARY;		PRT;	240 AA.	
AC	Q9IE32;					
DT	01-OCT-2000	(TrEMBLrel. 15, Created)				
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)				
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)				
DE	TM, GP41	(Fragment).				
GN	ENV.					
OS	Human immunodeficiency virus 1.					
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.					
NCBI_TaxID=11676;	[1]					
RN	SEQUENCE FROM N.A.					
RP	STRAIN=BCF11;					
RA	Roques P., Robertson D., Sousquiere S., Damond F., Mauchere P.,					
RA	Depienne C., Brun-Vezinet F., Dormont D., Simon F.O.;					
RT	Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";					
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.					
DR	EMBL:	AJ243365;	CAB96335.1;	-	-	
DR	GO:	GO:0016021;	C:integral to membrane;	IEA.		
DR	GO:	GO:0019031;	C:viral envelope;	IEA.		
DR	GO:	GO:0005198;	F:structural molecule activity;	IEA.		
DR	InterPro:	IPR000328;	Env_GP41.			
DR	Pfam:	PF00517;	GP41.1;			
DR	Transmembrane.					
FT	NON_TER	1				
FT	NON_TER	240				
SQ	SEQUENCE	240 AA;	27723 MW;	1E206E8D491AA197 CRC64;		
 Query Match						
Best Local Similarity 91.5%; Score 161; DB 15; Length 240;						
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;						
OY	1 ALETLNQQLLDLWGCRGLVCYTYSVRWNKT 32					
DB	:     :     :     :     :					
DB	50 AETLVNQQLLDLWGCRGLICYSVKWNKT 81					
 RESULT 7						
ID	Q9IHU8	PRELIMINARY;		PRT;	125 AA.	
AC	Q9IHU8;					
DT	01-OCT-2000	(TrEMBLrel. 15, Created)				
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)				
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)				
DE	Envelope glycoprotein (Fragment).					
GN	GP41.					
OS	Human immunodeficiency virus 1.					
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.					
NCBI_TaxID=11676;	[1]					
RN	SEQUENCE FROM N.A.					
RP	STRAIN=97CM766;					
RX	MEDLINE=20386754;	PubMed=10933623;				
RA	Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,					
RA	Pleniakz D., Schable C., Lal R.B.;					
RT	"Phylogenetic analysis of protease and transmembrane regions of HIV type 1 group O.";					





Db 29 ALETLLQNQQLLDLWGCKGRGLVCYTSVKWNKT 60

## RESULT 13

Q9IHV5 PRELIMINARY; PRT; 137 AA.  
AC Q9IHV5  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.

OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=97CM761;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,  
RT Pieniazek D., Schable C., Lal R.B.;  
RA "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O.";  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).

DR ENBL; AF229229; AAF71906.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 137  
SQ SEQUENCE 137 AA; 16494 MW; 55C47096D8168493 CRC64;

Query Match 90.3%; Score 159; DB 15; Length 137;  
Best Local Similarity 84.4%; Pred. No. 4.2e-16;  
Matches 27; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLLQNQQLLDLWGCKGRGLVCYTSVRWNET 32  
Db 30 ALETLLQNQQLLDLWGCKGRGLVCYTSVKWNKT 61

## RESULT 14

Q8J3Q4 PRELIMINARY; PRT; 155 AA.  
AC Q8J3Q4  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE GP41 protein (Fragment).  
GN ENV.

OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=22092513; PubMed=12097573;  
RA Boutonnet N., Janssens W., Bouton C., Verschelde J.L., Heyndrickx L.,  
RA Beirnaert E., van der Groen G., Laetens I.;  
RT "Comparison of predicted scaffold-compatible sequence variation in the  
triple-hairpin structure of human immunodeficiency virus type 1 gp41  
with patient data.";  
RL J. Virol. 76:7595-7606(2002).

DR ENBL; AJ427999; CAD20951.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 155

SQ SEQUENCE 155 AA; 18571 MW; 1EA7198FF9BB1F CRC64;

Query Match 90.3%; Score 159; DB 15; Length 155;  
Best Local Similarity 81.2%; Pred. No. 4.8e-16;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALETLLQNQQLLDLWGCKGRGLVCYTSVRWNET 32  
Db 38 ALETLLQNQQLLDLWGCKGRGLVCYTSVKWNKT 69

## RESULT 15

Q9IEA3 PRELIMINARY; PRT; 208 AA.  
AC Q9IEA3;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE GP41 (Fragment).  
GN ENV.

OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=YBF32;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RT Philippe M.;  
RL "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236417; CAB96265.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 24286 MW; 1C2961C1953A07A1 CRC64;

Query Match 90.3%; Score 159; DB 15; Length 208;  
Best Local Similarity 81.2%; Pred. No. 6.5e-16;  
Matches 26; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

SQ SEQUENCE 208 AA; 24286 MW; 1C2961C1953A07A1 CRC64;

QY 1 ALETLLQNQQLLDLWGCKGRGLVCYTSVRWNET 32  
Db 24 ALETLLQNQQLLDLWGCKGRGLVCYTSVRWNET 55

Search completed: May 7, 2004, 17:49:14  
Job time : 40.605 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:51:11 ; Search time 46.9552 Seconds  
(without alignments)  
132.382 Million cell updates/sec

Title: US-09-147-362A-14

Perfect score: 122  
Sequence: 1 LNQRLNSWGCKRLVCYTSV 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	22	2 AAW80472	Aw80472 Peptide d
2	122	100.0	28	2 AAW80473	Aw80473 Peptide d
3	105	86.1	33	3 AAB12231	Aab12231 Partial s
4	105	86.1	33	3 AAB12212	Aab12212 Partial s
5	105	86.1	40	2 AAW07346	Aw07346 Partial s
6	105	86.1	40	2 AAW07352	Aw07352 Partial s
7	104	85.2	32	2 AAW80469	Aw80469 Peptide d
8	103	84.4	33	3 AAB12236	Aab12236 Partial s
9	103	84.4	40	2 AAW07343	Aw07343 Partial s
10	99	81.1	33	3 AAB12210	Aab12210 Partial s
11	99	81.1	33	3 AAB12221	Aab12221 Partial s
12	99	81.1	33	3 AAB12213	Aab12213 Partial s
13	99	81.1	41	3 AAW07353	Aw07353 Partial s
14	99	81.1	42	2 AAW07350	Aw07350 Partial s
15	99	81.1	116	2 AAY05555	Aay05555 HIV-1 gro
16	99	81.1	117	2 AAY05548	Aay05548 HIV-1 gro
17	99	81.1	715	2 AAY05625	Aay05625 HIV-1 gro
18	98	80.3	32	2 AAW80470	Aw80470 Peptide d
19	98	80.3	33	3 AAB12211	Aab12211 Partial s
20	98	80.3	33	3 AAB12222	Aab12222 Partial s
21	98	80.3	33	3 AAB12235	Aab12235 Partial s
22	98	80.3	33	3 AAB12214	Aab12214 Partial s
23	98	80.3	33	3 AAB12220	Aab12220 Partial s
24	98	80.3	41	2 AAW07351	Aw07351 Partial s
25	98	80.3	113	2 AAY05559	Aay05559 HIV-1 gro

26	98	80.3	113	2 AAY05565	Aay05565 HIV-1 gro
27	98	80.3	115	2 AAY05557	Aay05557 HIV-1 gro
28	98	80.3	200	3 AAY77373	Aay77373 HIV-1 gro
29	98	80.3	215	2 AAY09499	Aay09499 HIV-1 gro
30	98	80.3	215	2 AAY06983	Recombina
31	98	80.3	215	3 AAY77374	Aay77374 HIV-1 gro
32	98	80.3	245	2 AAY09493	Aay09493 HIV-1 gro
33	98	80.3	245	2 AAY06977	Recombina
34	98	80.3	245	3 AAY77369	Aay77369 HIV-1 gro
35	98	80.3	281	2 AAY09507	Aay09507 HIV-1 gro
36	98	80.3	373	2 AAY09495	Aay09495 HIV-1 gro
37	98	80.3	373	2 AAY06979	Recombina
38	98	80.3	460	2 AAY09500	Aay09500 HIV-1 gro
39	98	80.3	460	2 AAY06984	Recombina
40	98	80.3	460	3 AAY77375	Aay77375 HIV-1 gro
41	98	80.3	474	3 AAY77371	Aay77371 HIV-1 gro
42	98	80.3	488	2 AAY09504	Aay09504 HIV-1 gro
43	98	80.3	490	2 AAY09494	Aay09494 HIV-1 gro
44	98	80.3	490	2 AAY06978	Recombina
45	98	80.3	490	3 AAY77370	Aay77370 HIV-1 gro

ALIGNMENTS

RESULT 1  
AAW80472  
ID AAW80472 standard; peptide; 22 AA.  
XX AC AAW80472;

XX XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
XX OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX XX  
PN WO9845323-A1.  
XX XX  
PD 15-OCT-1998.  
XX XX  
PF 06-APR-1998; 98WO-FR000691.  
XX XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX XX  
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX XX  
PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX XX  
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
XX XX  
PS Claim 6; Page 44; 55pp; French.  
XX XX  
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX XX  
SQ (Updated on 27-AUG-2003 to correct OS field.)

Sequence 22 AA;

Query Match 100.0%; Score 122; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 1.2e-10; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0;

QY 1 LNQQRLLNSWGCKGRLVCYTSV 22  
|||||  
Db 1 LNQQRLLNSWGCKGRLVCYTSV 22  
|||||

RESULT 2  
AAW80473  
ID AAW80473 standard; peptide; 28 AA.  
XX AC  
XX AAW80473;  
XX AC  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX Peptide derived from a conserved sequence of group O human HIV.  
XX Group O human immune deficiency virus; HIV; detection; infection.  
XX  
XX Synthetic.  
OS Human immunodeficiency virus 1.  
OS WO9845323-A1.  
PN  
XX  
PD 15-OCT-1998.  
XX  
XX 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
PA  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
PI  
XX WPI; 1998-583190/49.  
XX  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
PT  
XX  
PS Claim 6; Page 44; 55pp; French.  
XX  
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 28 AA;

Query Match 100.0%; Score 122; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.6e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNQQRLLNSWGCKGRLVCYTSV 22  
|||||  
Db 7 LNQQRLLNSWGCKGRLVCYTSV 28  
|||||

RESULT 3  
AAB12231  
ID AAB12231 standard; peptide; 33 AA.  
XX AC  
XX AAB12231;  
XX  
DT 12-SEP-2003 (revised)  
DT 10-NOV-2000 (first entry)  
XX  
XX

DE Partial sequence of HIV-1 strain MAN gp41 immunodominant region.  
XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
KW acquired immunodeficiency syndrome; group O HIV; gp41; MAN.  
XX  
XX Human immunodeficiency virus 1.  
XX  
XX EP1013766-A2.  
XX  
PD 28-JUN-2000.  
XX  
XX 29-NOV-1999; 99EP-00309491.  
XX  
PR 30-NOV-1998; 98US-0110292P.  
PR 08-FEB-1999; 99US-0119138P.  
PR 04-NOV-1999; 99US-00433428.  
XX  
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
PA  
XX De Leys R, Zheng J;  
PI  
XX WPI; 2000-402205/35.  
XX  
XX New antigenic peptides and peptide functional derivatives, useful for  
PT detection of antibodies produced in response to human immunodeficiency  
PT virus group O antibodies.  
XX  
XX Example 1; Fig 1; 52pp; English.  
XX  
CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
CC Virus Type 1 (HIV-1) strain MAN. HIV is the principle aetiological agent  
CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
CC protein, and may be used as an antigen for the detection of antibodies  
CC produced in response to HIV infection. MAN is a member of HIV group O  
CC (outlier). The present sequence is the immunodominant region of gp41.  
CC This sequence was used in a sequence homology alignment, which in turn  
CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
CC (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 33 AA;

Query Match 86.1%; Score 105; DB 3; Length 33;  
Best Local Similarity 81.8%; Pred. No. 5.4e-08;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLLNSWGCKGRLVCYTSV 22  
: |||||  
Db 9 IQNQQLLSWGCKGRLVCYTSV 30  
: |||||

RESULT 4  
AAB12212  
ID AAB12212 standard; peptide; 33 AA.  
XX AC  
XX AAB12212;  
XX  
DT 12-SEP-2003 (revised)  
DT 10-NOV-2000 (first entry)  
XX  
XX Partial sequence of HIV-1 strain BCF13 gp41 immunodominant region.  
XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF13.  
XX  
XX Human immunodeficiency virus 1.  
XX  
XX EP1013766-A2.  
XX  
PD 28-JUN-2000.  
XX  
XX 29-NOV-1999; 99EP-00309491.  
XX  
XX 30-NOV-1998; 98US-0110292P.  
PR

PT immunogens.

XX

XX Claim 12; Page 34; 71pp; French.

XX

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided

CC into 2 major groups based on the nucleotide sequences of the envelop gene

CC (env): group M containing sub-groups A-G, and group O containing the

CC strains AN770 and MVP180. The invention relates to the discovery of

CC several new strains of HIV-1 which can be placed in group O, based on the

CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39

CC and AAM07329-64). The novel strains have been deposited as retroviruses

CC CINC I-1544 (BCF02 (ESG)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547

CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from

CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein

CC encoded by the env gene. The nucleic acids can be used to detect gp. O

CC HIV-1 strains by hybridisation or (as primers) by gene amplification.

CC also for screening and typing of such strains. Peptides encoded by the

CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O

CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-

CC 2003 to standardise OS field)

XX

XX Sequence 40 AA;

XX

Query Match 86.1%; Score 105; DB 2; Length 40;

Best Local Similarity 81.8%; Pred. No. 6.6e-08;

Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0

QY 1 LNQQRLNSWGCKGRLVCYTSV 22

DB 14 IQNQQRLNSWGCKGRLVCYTSV 35

1 : |||||

14 : |||||

RESULT 6

AAW07352

ID AAW07352 standard; peptide; 40 AA.

AC AAW07352;

XX

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF13.

XX

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;

KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;

KX primer; hybridisation; amplification; PCR; polymerase chain reaction;

XX immunogen; antibody.

XX

XX Human immunodeficiency virus 1.

OS

OS W09627013-A1.

PN

XX

XX 06-SEP-1996.

PD

XX

XX 26-FEB-1996; 96WO-FR000294.

XX

XX 27-FEB-1995; 95FR-00002236.

PR

XX

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PA

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX

XX Simon F, Saragosti S, Louseertajaka I, Ly T, Chaixbaudier M;

PI

XX

XX WPI; 1996-412779/41.

DR

XX

XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and

PT antibodies - useful for diagnosis, screening and typing, or as

PT immunogens.

XX

XX Claim 12; Page 46; 71pp; French.

PS

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided

CC



```

AC AAW07343;
XX
XX 16-OCT-2003 (revised)
DT
XX 25-MAR-2003 (revised)
DT
XX 03-JUN-1997 (first entry)
DT
XX
XX Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (ESS).
DE
XX
XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
KW primer; hybridisation; amplification; PCR; polymerase chain reaction;
KW immunogen; antibody.
XX
XX Human immunodeficiency virus 1.
OS
XX
XX WO9627013-A1.
PN
XX
XX 06-SEP-1996.
PD
XX
XX 26-FEB-1996; 96WO-FR000294.
PF
XX
XX 27-FEB-1995; 95FR-00002236.
PR
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA
XX Simon F, Saragosti S, Lousseertajaka I, Ly T, Chaixbaudier M;
PI
XX
XX WPI; 1996-412779/41.
DR
XX N-PSDB; AAT44918.
DR
XX
XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and
PT antibodies - useful for diagnosis, screening and typing, or as
PT immunogens.
PT
XX
XX Claim 12; Page 33; 71pp; French.
PS
XX
XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
CC into 2 major groups based on the nucleotide sequences of the envelop gene
CC (env): group M containing sub-groups A-G, and group O containing the
CC strains ANT70 and MVP5180. The invention relates to the discovery of
CC several new strains of HIV-1 which can be placed in group O, based on the
CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
CC and AAW07329-64). The novel strains have been deposited as retroviruses
CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
CC the strain BCF02 (ESS) and corresponds to a fragment of the gp41 protein
CC encoded by the env gene. The nucleic acids can be used to detect gp. O
CC HIV-1 strains by hybridisation or (as primers) by gene amplification,
CC also for screening and typing of such strains. Peptides encoded by the
CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
CC 2003 to standardise OS field)
XX
XX
XX Sequence 40 AA;
SQ
Query Match 84.4%; Score 103; DB 2; Length 40;
Best Local Similarity 77.3%; Pred. No. 1.3e-07;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCYTSV 22
: |:|||||:|||||:
Db 14 IQNQQLNLWGCKGRVLCYTSV 35

RESULT 10
AAB12210
ID AAB12210 standard; peptide; 33 AA.
XX
XX AAB12210;
AC
XX
XX 12-SEP-2003 (revised)
DT
XX 10-NOV-2000 (first entry)
DT

Partial sequence of HIV-1 strain ABT2156 gp41 immunodominant region.
DE
XX
XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF09.
KW
XX
XX Human immunodeficiency virus 1.
OS
XX
XX EP1013766-A2.
PN
XX
XX 28-JUN-2000.
PD
XX
XX 29-NOV-1999; 99EP-00309491.
PF
XX
XX 30-NOV-1998; 98US-0110292P.
PR
XX 08-FEB-1999; 99US-0119138P.
PR
XX 04-NOV-1999; 99US-00433428.
PR
XX
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
PA
XX
XX De Leys R, Zheng J;
XX
XX WPI; 2000-402205/35.
DR
XX
XX New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
PT
XX
XX Example 1; Fig 1; 52pp; English.
PS
XX
XX The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain BCF09. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. BCF09 is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment, which
CC in turn was used to derive a consensus sequence peptide: peptide 147
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
XX
XX Sequence 33 AA;
SQ
Query Match 81.1%; Score 99; DB 3; Length 33;
Best Local Similarity 77.3%; Pred. No. 4e-07;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCYTSV 22
: |:|||||:|||||:
Db 9 IQNQQLNLWGCKGRVLCYTSV 30

RESULT 11
AAB12221
ID AAB12221 standard; peptide; 33 AA.
XX
XX AAB12221;
AC
XX
XX 12-SEP-2003 (revised)
DT
XX 10-NOV-2000 (first entry)
DT

Partial sequence of HIV-1 strain ABT2156 gp41 immunodominant region.
DE
XX
XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; ABT2156.
KW
XX
XX Human immunodeficiency virus 1.
OS
XX
XX EP1013766-A2.
PN
XX
XX 28-JUN-2000.
PD
XX
XX 29-NOV-1999; 99EP-00309491.
PF
XX
XX

```

```

PR 30-NOV-1998; 98US-0110292P.
PR 08-FEB-1999; 99US-0119138P.
PR 04-NOV-1999; 99US-00433428.
XX
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
XX De Leys R, Zheng J;
XX
XX WPI; 2000-402205/35.
XX
XX New antigenic peptides and peptide functional derivatives, useful for
XX PT detection of antibodies produced in response to human immunodeficiency
XX PT virus group O antibodies.
XX
XX
XX Example 1; Fig 1; 52pp; English.
XX
XX The present sequence is a partial gp41 protein of Human Immunodeficiency
XX CC Virus Type 1 (HIV-1) strain ABB2156. HIV is the principle aetiological
XX CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
XX CC envelope protein, and may be used as an antigen for the detection of
XX CC antibodies produced in response to HIV infection. ABB2156 is a member of
XX CC HIV group O (outlier). The present sequence is the immunodominant region
XX CC of gp41. This sequence was used in a sequence homology alignment, which
XX CC in turn was used to derive a consensus sequence peptide; peptide 147
XX CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 33 AA;
XX
XX Query Match 81.1%; Score 99; DB 3; Length 33;
XX Best Local Similarity 77.3%; Pred. No. 4e-07;
XX Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 LNQQRLNSWGCKGRVCYTSV 22
XX :|:|||||
XX Db 9 IQNQQLLNSWGCKGRVCYTSV 30
XX
XX RESULT 12
XX AAB12213
XX ID AAB12213 standard; peptide; 33 AA.
XX AC AAB12213;
XX
XX DT 12-SEP-2003 (revised)
XX DT 10-NOV-2000 (first entry)
XX
XX DE Partial sequence of HIV-1 strain BCF14 gp41 immunodominant region.
XX
XX KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
XX KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF14.
XX
XX OS Human immunodeficiency virus 1.
XX
XX XX EP1013766-A2.
XX
XX PD 28-JUN-2000.
XX
XX PF 29-NOV-1999; 99EP-00309491.
XX
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
XX De Leys R, Zheng J;
XX
XX WPI; 2000-402205/35.
XX
XX New antigenic peptides and peptide functional derivatives, useful for
XX PT detection of antibodies produced in response to human immunodeficiency
XX PT virus group O antibodies.
XX
XX
XX Example 1; Fig 1; 52pp; English.
XX
XX The present sequence is a partial gp41 protein of Human Immunodeficiency
XX CC Virus Type 1 (HIV-1) strain ABB2156. HIV is the principle aetiological
XX CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
XX CC envelope protein, and may be used as an antigen for the detection of
XX CC antibodies produced in response to HIV infection. ABB2156 is a member of
XX CC HIV group O (outlier). The present sequence is the immunodominant region
XX CC of gp41. This sequence was used in a sequence homology alignment, which
XX CC in turn was used to derive a consensus sequence peptide; peptide 147
XX CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 33 AA;
XX
XX Query Match 81.1%; Score 99; DB 3; Length 33;
XX Best Local Similarity 77.3%; Pred. No. 4e-07;
XX Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 LNQQRLNSWGCKGRVCYTSV 22
XX :|:|||||
XX Db 9 IQNQQLLNSWGCKGRVCYTSV 30
XX
XX RESULT 13
XX AAW07353
XX ID AAW07353 standard; peptide; 41 AA.
XX
XX AC AAW07353;
XX
XX DT 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 03-JUN-1997 (first entry)
XX
XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF14.
XX
XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
XX KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
XX KW primer; hybridisation; amplification; PCR; polymerase chain reaction;
XX KW immunogen; antibody.
XX
XX OS Human immunodeficiency virus 1.
XX
XX XX WO9627013-A1.
XX
XX PD 06-SEP-1996.
XX
XX PF 26-FEB-1996; 96WO-FR000294.
XX
XX PR 27-FEB-1995; 95FR-00002236.
XX
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX
XX PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;
XX
XX WPI; 1996-412779/41.
XX
XX PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and
XX PT antibodies - useful for diagnosis, screening and typing, or as
XX PT immunogens.
XX
XX PS Claim 12; Page 46; 71pp; French.
XX
XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
XX CC into 2 major groups based on the nucleotide sequences of the envelop gene
XX CC (env): group M containing sub-groups A-G, and group O containing the
XX CC strains ANT70 and MVP5180. The invention relates to the discovery of
XX CC several new strains of HIV-1 which can be placed in group O, based on the
XX CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
XX CC and AAW07329-64). The novel strains have been deposited as retroviruses
XX CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
XX CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
XX CC the strain BCF14 and corresponds to a fragment of the gp41 protein

```



Human immunodeficiency virus type 1 (HIV-1) strains are currently divided into 2 major groups based on the nucleotide sequences of the envelop gene (env): group M containing sub-groups A-G, and group O containing the strains AN770 and MVP5180. The invention relates to the discovery of several new strains of HIV-1 which can be placed in group O, based on the partial sequences of the C235-env, gp41 and gag genes (see AN744907-39 and AAM07329-64). The novel strains have been deposited as retroviruses NCIM I-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547 (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from the strain BCF09 and corresponds to a fragment of the gp41 protein encoded by the env gene. The nucleic acids can be used to detect gp. O HIV-1 strains by hybridisation or (as primers) by gene amplification, also for screening and typing of such strains. Peptides encoded by the nucleic acids can be used as immunogens to raise Ab for detecting gp. O HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-2004)

Search completed: May 7, 2004, 18:31:55  
Job time : 46.9552 secs

---

Result No.	Score	Query Match	Length	DB	ID	Description
1	105	86.1	33	3	US-09-433-428D-6	Sequence 6, Appl
2	105	86.1	33	3	US-09-433-428D-25	Sequence 25, Appl
3	105	86.1	40	3	US-08-894-699-39	Sequence 39, Appl
4	105	86.1	40	3	US-08-894-699-68	Sequence 68, Appl
5	105	86.1	40	3	US-09-444-410-39	Sequence 39, Appl
6	105	86.1	40	3	US-09-444-410-68	Sequence 68, Appl
7	103	84.4	33	3	US-09-433-428D-30	Sequence 30, Appl
8	103	84.4	33	3	US-08-894-699-36	Sequence 36, Appl
9	103	84.4	40	3	US-09-444-410-36	Sequence 36, Appl
10	99	81.1	33	3	US-09-433-428D-4	Sequence 4, Appl
11	99	81.1	33	3	US-09-433-428D-7	Sequence 7, Appl
12	99	81.1	33	3	US-09-433-428D-15	Sequence 15, Appl
13	99	81.1	41	3	US-08-894-699-69	Sequence 69, Appl
14	99	81.1	41	3	US-09-444-410-69	Sequence 69, Appl
15	99	81.1	42	3	US-08-894-699-66	Sequence 66, Appl
16	99	81.1	42	3	US-09-444-410-66	Sequence 66, Appl
17	99	81.1	116	4	US-09-462-917A-20	Sequence 20, Appl
18	99	81.1	117	4	US-09-462-917A-6	Sequence 6, Appl
19	99	81.1	715	3	US-09-462-917A-134	Sequence 134, App
20	98	80.3	33	3	US-09-433-428D-5	Sequence 5, Appl
21	98	80.3	33	3	US-09-433-428D-8	Sequence 8, Appl
22	98	80.3	33	3	US-09-433-428D-14	Sequence 14, Appl
23	98	80.3	33	3	US-09-433-428D-16	Sequence 16, Appl
24	98	80.3	33	3	US-09-433-428D-29	Sequence 29, Appl
25	98	80.3	41	3	US-08-894-699-67	Sequence 67, Appl
26	98	80.3	41	3	US-09-444-410-67	Sequence 67, Appl
27	98	80.3	113	4	US-09-462-917A-28	Sequence 28, Appl



```
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-444-410-39

Query Match      86.1%; Score 105; DB 3; Length 40;
Best Local Similarity 81.8%; Pred. No. 5.2e-09;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 LNQORLLNSWGCKGRIVCYTSV 22
Db      14 IQNQQLLSWGCKGRIVCYTSV 35

RESULT 6
US-09-444-410-68
; Sequence 68, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-444-410-68

Query Match      86.1%; Score 105; DB 3; Length 40;
Best Local Similarity 81.8%; Pred. No. 5.2e-09;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 LNQORLLNSWGCKGRIVCYTSV 22
Db      14 IQNQQLLSWGCKGRIVCYTSV 35

RESULT 7
US-09-433-428D-30
; Sequence 30, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; US-09-433-428D-30

Query Match      84.4%; Score 103; DB 3; Length 33;
Best Local Similarity 77.3%; Pred. No. 8.5e-09;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 LNQORLLNSWGCKGRIVCYTSV 22
Db      9 IQNQQLLSWGCKGRIVCYTSV 30

RESULT 8
US-08-894-699-36
; Sequence 36, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
```

/ TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
/ TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
/ NUMBER OF SEQUENCES: 81  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
/ ADDRESSEE: P.C.  
/ STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
/ STREET: FLOOR  
/ CITY: ARLINGTON  
/ STATE: VA  
/ COUNTRY: USA  
/ ZIP: 22202  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/894,699  
/ FILING DATE: 01-DEC-1997  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/FR96/00294  
/ FILING DATE: 26-FEB-1996  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: FR 95/02236  
/ FILING DATE: 27-FEB-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: OBLON, NORMAN F.  
/ REGISTRATION NUMBER: 24,614  
/ REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 703-413-3000  
/ TELEFAX: 703-413-2220  
/ INFORMATION FOR SEQ ID NO: 36:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 40 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-08-894-699-36

Query Match 84.4%; Score 103; DB 3; Length 40;  
Best Local Similarity 77.3%; Pred. No. 1e-08;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRIVCYTSV 22  
Db 14 IQNQQLNSWGCKGRIVCYTSV 35

RESULT 9  
US-09-444-410-36  
; Sequence 36, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSHI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
/ TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
/ NUMBER OF SEQUENCES: 81  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
/ ADDRESSEE: P.C.  
/ STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
/ STREET: FLOOR  
/ CITY: ARLINGTON  
/ STATE: VA  
/ COUNTRY: USA

/ ZIP: 22202  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/444,410  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/894,699  
/ FILING DATE:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: FR 95/02236  
/ FILING DATE: 27-FEB-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: OBLON, NORMAN F.  
/ REGISTRATION NUMBER: 24,614  
/ REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 703-413-3000  
/ TELEFAX: 703-413-2220  
/ INFORMATION FOR SEQ ID NO: 36:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 40 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
/ US-09-444-410-36

Query Match 84.4%; Score 103; DB 3; Length 40;  
Best Local Similarity 77.3%; Pred. No. 1e-08;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRIVCYTSV 22  
Db 14 IQNQQLNSWGCKGRIVCYTSV 35

RESULT 10  
US-09-433-428D-4  
; Sequence 4, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 33  
; TYPE: PPT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-4

Query Match 81.1%; Score 99; DB 3; Length 33;  
Best Local Similarity 77.3%; Pred. No. 3.4e-08;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRIVCYTSV 22  
Db 9 IQNQQLNSWGCKGRIVCYTSV 30

RESULT 11  
US-09-433-428D-7  
; Sequence 7, Application US/09433428D  
; Patent No. 6149910

```
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-7

Query Match      81.1%; Score 99; DB 3; Length 33;
Best Local Similarity 77.3%; Pred. No. 3.4e-08;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCVTSV 22
   : ||||| ||||| |||||
Db 9 IQNQQLLSWGCKGRQVCYTSV 30

RESULT 12
US-09-433-428D-15
; Sequence 15, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-15

Query Match      81.1%; Score 99; DB 3; Length 33;
Best Local Similarity 77.3%; Pred. No. 3.4e-08;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCVTSV 22
   : ||||| ||||| |||||
Db 9 IQNQQLLSWGCKGRVLCVTSV 30

RESULT 13
US-08-894-699-69
; Sequence 69, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
```

```
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-699-69

Query Match      81.1%; Score 99; DB 3; Length 41;
Best Local Similarity 77.3%; Pred. No. 4.2e-08;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCVTSV 22
   : ||||| ||||| |||||
Db 14 IQNQQLLSWGCKGRQVCYTSV 35

RESULT 14
US-09-444-410-69
; Sequence 69, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-69

Query Match 81.1%; Score 99; DB 3; Length 41;  
Best Local Similarity 77.3%; Pred. No. 4.2e-08;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LNQQRLNSWGCKGRLVCYTSV 22  
Db 14 IQNQQLNSWGCKGRQVCYTSV 35

## RESULT 15

US-08-894-699-66  
Sequence 66, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DIONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSER: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,699  
FILING DATE: 01-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR96/00294  
FILING DATE: 26-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-699-66

Query Match 81.1%; Score 99; DB 3; Length 42;  
Best Local Similarity 77.3%; Pred. No. 4.3e-08;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LNQQRLNSWGCKGRLVCYTSV 22  
Db 14 IQNQQLNSWGCKGRLVCYTSV 35

Search completed: May 7, 2004, 18:36:06  
Job time : 13.791 secs



Result No.	Score	Query Match	Length	DB	ID	Description
1	99	81.1	116	14	US-10-320-786-20	Sequence 20, Appl
2	99	81.1	117	14	US-10-320-786-6	Sequence 6, Appl
3	99	81.1	715	14	US-10-320-786-134	Sequence 134, Appl
4	98	80.3	113	14	US-10-320-786-28	Sequence 28, Appl
5	98	80.3	113	14	US-10-320-786-40	Sequence 40, Appl
6	98	80.3	115	14	US-10-320-786-24	Sequence 24, Appl
7	98	80.3	215	8	US-08-911-824-58	Sequence 58, Appl
8	98	80.3	245	8	US-08-911-824-48	Sequence 48, Appl
9	98	80.3	281	8	US-08-911-824-120	Sequence 120, Appl
10	98	80.3	373	8	US-08-911-824-52	Sequence 52, Appl
11	98	80.3	460	8	US-08-911-824-60	Sequence 60, Appl
12	98	80.3	488	8	US-08-911-824-95	Sequence 95, Appl
13	98	80.3	490	8	US-08-911-824-50	Sequence 50, Appl
14	98	80.3	526	8	US-08-911-824-97	Sequence 97, Appl
15	98	80.3	618	8	US-08-911-824-54	Sequence 54, Appl

```
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-6

Query Match      81.1%; Score 99; DB 14; Length 117;
Best Local Similarity 77.3%; Pred. No. 1.3e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db 34 IQNQQLNLWGCKGRLVCYTSV 55

RESULT 3
US-10-320-786-134
; Sequence 134, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-134

Query Match      81.1%; Score 99; DB 14; Length 715;
Best Local Similarity 77.3%; Pred. No. 7.2e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db 623 IQNQQLNLWGCKGRLVCYTSV 644

RESULT 4
US-10-320-786-28
; Sequence 28, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
```

```
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-28

Query Match      80.3%; Score 98; DB 14; Length 113;
Best Local Similarity 72.7%; Pred. No. 1.8e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db 34 IQNQQLNLWGCKGRLVCYTSV 55

RESULT 5
US-10-320-786-40
; Sequence 40, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: Misc_feature
; LOCATION: (1)..(113)
; OTHER INFORMATION: xaa = unknown
US-10-320-786-40

Query Match      80.3%; Score 98; DB 14; Length 113;
Best Local Similarity 72.7%; Pred. No. 1.8e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db 34 IRNQQLNLWGCKGRLVCYTSV 55

RESULT 6
US-10-320-786-24
; Sequence 24, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
```

; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 24  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-24

Query Match 80.3%; Score 98; DB 14; Length 115;  
Best Local Similarity 72.7%; Pred. No. 1.8e-06;  
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LNQQRLLNSWGCKGRILVCYTSV 22  
: ||| |||||: |||||  
Db 34 IQSQQLNLNLWGCKGRILVCYTSV 55

RESULT 7  
US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 80.3%; Score 98; DB 8; Length 215;  
Best Local Similarity 72.7%; Pred. No. 3.2e-06;  
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LNQQRLLNSWGCKGRILVCYTSV 22  
: ||| |||||: |||||  
Db 122 IQNQQLNLNLWGCKGRILVCYTSV 143

RESULT 8  
US-08-911-824-48  
; Sequence 48, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE

; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 48  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL  
US-08-911-824-48

Query Match 80.3%; Score 98; DB 8; Length 245;  
Best Local Similarity 72.7%; Pred. No. 3.7e-06;  
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LNQQRLLNSWGCKGRILVCYTSV 22  
: ||| |||||: |||||  
Db 122 IQNQQLNLNLWGCKGRILVCYTSV 143

RESULT 9  
US-08-911-824-120  
; Sequence 120, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 120  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL  
US-08-911-824-120

Query Match 80.3%; Score 98; DB 8; Length 281;  
Best Local Similarity 72.7%; Pred. No. 4.2e-06;  
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LNQQRLLNSWGCKGRILVCYTSV 22  
: ||| |||||: |||||  
Db 122 IQNQQLNLNLWGCKGRILVCYTSV 143

RESULT 10  
US-08-911-824-52  
; Sequence 52, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

```
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL
; US-08-911-824-52
```

```
Query Match      80.3%; Score 98; DB 8; Length 373;
Best Local Similarity 72.7%; Pred. No. 5.4e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LNQQRLNSWGCKGRLCVTSV 22
   : |::| |::| |::| |::|
Db 122 IQNQQLLNLWGCKGRLCVTSV 143
```

```
RESULT 11
US-08-911-824-60
; Sequence 60, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS
; US-08-911-824-60
```

```
Query Match      80.3%; Score 98; DB 8; Length 460;
Best Local Similarity 72.7%; Pred. No. 6.6e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LNQQRLNSWGCKGRLCVTSV 22
   : |::| |::| |::| |::|
Db 367 IQNQQLLNLWGCKGRLCVTSV 388
```

```
RESULT 12
US-08-911-824-95
; Sequence 95, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911.824
```

```
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-14PL
; US-08-911-824-95
```

```
Query Match      80.3%; Score 98; DB 8; Length 488;
Best Local Similarity 72.7%; Pred. No. 7e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LNQQRLNSWGCKGRLCVTSV 22
   : |::| |::| |::| |::|
Db 122 IQNQQLLNLWGCKGRLCVTSV 143
```

```
RESULT 13
US-08-911-824-50
; Sequence 50, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911.824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
; US-08-911-824-50
```

```
Query Match      80.3%; Score 98; DB 8; Length 490;
Best Local Similarity 72.7%; Pred. No. 7e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LNQQRLNSWGCKGRLCVTSV 22
   : |::| |::| |::| |::|
Db 367 IQNQQLLNLWGCKGRLCVTSV 388
```

```
RESULT 14
US-08-911-824-97
; Sequence 97, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911.824
```

```

; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS
US-08-911-824-97

```

```

Query Match      80.3%; Score 98; DB 8; Length 526;
Best Local Similarity 72.7%; Pred. No. 7.5e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 LNQORLLNSWGCKGRGLVCYTSV 22
       : ||||| ||||| ||||| |||||
Db      367 IQNQQLLNLMGCKGRGLICYTSV 388

```

```

RESULT 15
US-08-911-824-54
; Sequence 54, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165,US,01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

```

```

Query Match      80.3%; Score 98; DB 8; Length 618;
Best Local Similarity 72.7%; Pred. No. 8.8e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 LNQORLLNSWGCKGRGLVCYTSV 22
       : ||||| ||||| ||||| |||||
Db      367 IQNQQLLNLMGCKGRGLICYTSV 388

```

Search completed: May 7, 2004, 18:47:36  
Job time : 35.1343 secs

Casey XV/11 EOL/2 S.H.

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	93	76.2	863	2	A53034	gag polyprotein -	
2	92	75.4	877	2	S49197	envlope protein p	
3	87	71.3	104	2	S52930	GP41 ENV protein -	
4	77	63.1	357	2	S21990	envlope protein g	
5	77	63.1	358	2	S22002	envlope protein g	
6	77	63.1	854	1	VCLJ51	env polyprotein pr	
7	76	62.3	358	2	S22000	envlope protein g	
8	76	62.3	358	2	S70417	envlope protein g	
9	76	62.3	855	2	A45713	Env transmembrane	
10	74	60.7	357	2	S22006	envlope protein g	
11	74	60.7	357	2	S21994	envlope protein g	
12	74	60.7	357	2	S22004	envlope protein g	
13	74	60.7	357	2	S21996	envlope protein g	
14	74	60.7	357	2	S21992	envlope protein g	
15	74	60.7	358	2	S21998	envlope protein g	
16	74	60.7	443	2	C41621	env polyprotein p	
17	74	60.7	445	2	A41621	env polyprotein M	
18	74	60.7	454	2	B41621	env polyprotein pr	
19	74	60.7	843	1	H44001	env polyprotein pr	
20	74	60.7	847	2	T09448	envlope glycoprot	
21	74	60.7	847	2	S13289	env protein - huma	
22	74	60.7	852	1	VCLJBR	env polyprotein -	
23	74	60.7	852	2	T12016	envlope glycoprot	
24	74	60.7	853	2	S54384	envlope polyprot	
25	74	60.7	854	2	S13288	env protein - huma	
26	74	60.7	855	1	VCLJAZ	env polyprotein pr	
27	74	60.7	855	1	VCLJZR	env polyprotein pr	
28	74	60.7	856	1	VCLJH3	env polyprotein pr	
29	74	60.7	856	1	VCLJVL	env polyprotein pr	

Query Match 75.4%; Score 92; DB 2; Length 877;

```
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 602 IQNQQLLNWGCGRKRLICVTSV 623

RESULT 3
S22930
G:41 ENV protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Jun-1995 #sequence_revision 21-Jul-1995 #text_change 26-Aug-1999
C:Accession: S52930
R:Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de
submitted to the EMBL Data Library, January 1995
A:Description: A novel HIV-1 O strain illustrates the diversity of the O group.
A:Reference number: S52929
A:Accession: S52930
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-104 <COH>
A:Cross-references: EMBL:X84328; NID:G95526; PIDN:CAA59066.1; PID:G95527
C:Superfamily: type E retrovirus env polyprotein

Query Match 71.3%; Score 87; DB 2; Length 104;
Best Local Similarity 59.1%; Pred. No. 5,1e-06;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 40 MQNQQLLNWGCGRKAICVTSV 61

RESULT 4
S21990
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: simian immunodeficiency virus SIVcpz
A:Note: host Pan troglodytes (chimpanzee)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09990
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61357; NID:G60175; PIDN:CAA43626.1; PID:G60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
C:Superfamily: type E retrovirus env polyprotein

Query Match 63.1%; Score 77; DB 2; Length 357;
Best Local Similarity 59.1%; Pred. No. 0.00051;
Matches 13; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 88 LKQQLLGIWGCGRKRLICTTAV 109

RESULT 5
S22002
env polyprotein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 3L

Query Match 63.1%; Score 77; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 0.0011;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 577 LQDQQLGIWGCGRKAVCYTTV 598

RESULT 7
S22000
env polyprotein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
```

```
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <STE1>
A:Cross-references: EMBL:X61352; NID:G60186; PIDN:CAA43616.1; PID:G60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333,'X',335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:G60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 63.1%; Score 77; DB 2; Length 358;
Best Local Similarity 59.1%; Pred. No. 0.00051;
Matches 13; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 89 LKQQLLGIWGCGRKRLICTTAV 110

RESULT 6
VCLJSI
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: simian immunodeficiency virus SIVcpz
A:Note: host Pan troglodytes (chimpanzee)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09990
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
submitted to the EMBL Data Library, July 1991
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09990
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-854 <HUE>
A:Cross-references: EMBL:X52154; NID:G58866; PIDN:CAA36407.1; PID:G58874
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-500/Product: coat protein gp120 #status predicted <CP1>
F:501-854/Product: coat protein gp41 #status predicted <CP2>
F:501-517/Domain: transmembrane #status predicted <TM1>
F:675-693/Domain: transmembrane #status predicted <TM2>
F:805-821/Domain: transmembrane #status predicted <TM3>
F:134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,455

Query Match 63.1%; Score 77; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 0.0011;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
      :|:|:| | | | | | | | | | | | | | |
Db 577 LQDQQLGIWGCGRKAVCYTTV 598

RESULT 7
S22000
env polyprotein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
```



Db 589 LKQQLLNSWGCAFRQVCHTTV 610

RESULT 10

S22006

envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S70420; S22006

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by HIV-1 in blood and cerebrospinal fluid as determined

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70420

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-357 <ST2>

A:Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191

A:Experimental source: patient L

A:Note: submitted to the EMBL Data Library, July 1991

C:Superfamily: type E retrovirus env polyprotein

Query Match 60.7%; Score 74; DB 2; Length 357;

Best Local Similarity 54.5%; Pred. No. 0.0014;

Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNOQRLNSWGCKGRLVCYTSV 22

Db 88 LKQQLLGIWGCGRKLCITTV 109

RESULT 11

S21994

envelope protein gp120/gp41 - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 27B

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000

C:Accession: S21994; S70421

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by HIV-1 in blood and cerebrospinal fluid as determined

A:Reference number: S21990

A:Accession: S21994

A:Molecule type: DNA

A:Residues: 1-357 <STEL>

A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by HIV-1 in blood and cerebrospinal fluid as determined

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70421

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-140, 'X', 142-312, 'X', 314-357 <ST2>

A:Cross-references: EMBL:X61355; NID:g60179

C:Superfamily: type E retrovirus env polyprotein

Query Match 60.7%; Score 74; DB 2; Length 357;

Best Local Similarity 54.5%; Pred. No. 0.0014;

Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNOQRLNSWGCKGRLVCYTSV 22

Db 88 LKQQLLGIWGCGRKLCITTV 109

RESULT 12

S22004

envelope protein gp120/gp41 - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 4B

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000

C:Accession: S22004; S70419



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:53:36 ; Search time 7.8806 Seconds  
(without alignments)  
145.363 Million cell updates/sec

Title: US-09-147-362A-14  
Perfect score: 122  
Sequence: 1 LMQORLLNSWGCKRLVYCVTSV 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	77	63.1	854	1 ENV_SIVCZ	P17281 chimpanzee
2	75	61.5	857	1 ENV_HV2KA	Q74126 human immun
3	74	60.7	843	1 ENV_HV1Y2	P35961 human immun
4	74	60.7	847	1 ENV_HV1S1	P19550 human immun
5	74	60.7	847	1 ENV_HV1W2	P05880 human immun
6	74	60.7	851	1 ENV_HV1B8	P04582 human immun
7	74	60.7	852	1 ENV_HV1B9	P12488 human immun
8	74	60.7	852	1 ENV_HV1S3	P19549 human immun
9	74	60.7	853	1 ENV_HV1MF	P19551 human immun
10	74	60.7	853	1 ENV_HV1D2	P12487 human immun
11	74	60.7	855	1 ENV_HV1A2	P03378 human immun
12	74	60.7	855	1 ENV_HV1OY	P20888 human immun
13	74	60.7	855	1 ENV_HV1Z6	P04580 human immun
14	74	60.7	856	1 ENV_HV1B1	P03375 human immun
15	74	60.7	856	1 ENV_HV1H2	Q70626 human immun
16	74	60.7	856	1 ENV_HV1LW	P04578 human immun
17	74	60.7	856	1 ENV_HV1MN	P05877 human immun
18	74	60.7	856	1 ENV_HV1PV	P03376 human immun
19	74	60.7	856	1 ENV_HV1SC	P05878 human immun
20	74	60.7	856	1 ENV_HV1W1	P31872 human immun
21	74	60.7	861	1 ENV_HV1BR	P03377 human immun
22	74	60.7	865	1 ENV_HV1RH	P04579 human immun
23	74	60.7	867	1 ENV_HV1J3	P12489 human immun
24	74	60.7	868	1 ENV_HV1C4	P05879 human immun
25	73	59.8	856	1 ENV_HV2NZ	P05883 human immun
26	72	59.0	848	1 ENV_HV1JR	P20871 human immun
27	72	59.0	856	1 ENV_HV1H3	P04624 human immun
28	71	58.2	856	1 ENV_HV1ZH	P05881 human immun
29	70	57.4	712	1 ENV_HV2S2	P32536 human immun
30	70	57.4	846	1 ENV_HV1ND	P18799 human immun
31	70	57.4	851	1 ENV_HV2D1	P17755 human immun
32	70	57.4	851	1 ENV_HV2G1	P18040 human immun
33	70	57.4	858	1 ENV_HV2RO	P04577 human immun

RESULT 1				
ENV_SIVCZ	1	57.4	859	1 ENV_HV2CA
ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroviridae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."			
RL	Nature 345:356-359(1990).			
CC	-1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X52154; CAA36407.1; ..			
DR	PIR; S09990; VCLJSI.			
DR	HIV; X52154; ENVSCPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. .) (POTENTIAL).

ALIGNMENTS

```

FT CARBOHYD 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 63.1%; Score 77; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 0.00018;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 577 LQDQQLGLWGCGRAVCYTTV 598

```

```

RESULT 2
ENV_HV2KR
ID ENV_HV2KR STANDARD; PRT; 857 AA.
AC Q74126;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate KR) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=73484;
RN [1]
RP SEQUENCE FROM N.A.
RA Kraus G.K., Talbott R., Leavitt M., Luznick L., Schmidt A.,
RA Badel P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U22047; AAA64582.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 503
FT CHAIN 504 857
FT CARBOHYD 36 36
FT CARBOHYD 69 69
FT CARBOHYD 78 78
FT CARBOHYD 113 113
FT CARBOHYD 121 121
FT CARBOHYD 134 134
FT CARBOHYD 142 142
FT CARBOHYD 159 159
FT CARBOHYD 186 186
FT CARBOHYD 198 198
FT CARBOHYD 230 230
FT CARBOHYD 230 230

```

```

FT CARBOHYD 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 281 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 857 AA; 98689 MW; F881C675B85746DF CRC64;

Query Match 61.5%; Score 75; DB 1; Length 857;
Best Local Similarity 63.6%; Pred. No. 0.00038;
Matches 14; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 578 LRDQARLNSWGCAFRQVCYTTV 599

RESULT 3
ENV_HV1Y2
ID ENV_HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR FIR; H44001; H44001.
DR PDB; 1G9N; 27-DEC-00.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 489
FT CHAIN 490 843
FT TRANSMEM 738 755
FT DISULFID 53 73
FT DISULFID 118 201
FT DISULFID 125 192
FT DISULFID 130 155

```

```
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 60.7%; Score 74; DB 1; Length 843;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCVTSV 22
Db 574 LRDQQLLGWCGSKGLICTTAV 595

RESULT 4
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; M65024; AAA45072.1; -.
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENV$SF162.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7FP2AB CRC64;

Query Match 60.7%; Score 74; DB 1; Length 847;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQORLLNSWGCKGRVLCVTSV 22
Db 578 LKDQQLLGWCGSKGLICTTAV 599

RESULT 5
ENV_HV1W2
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
```



```
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 60.7%; Score 74; DB 1; Length 851;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNOQRLLNSWGCKGRIVCVTSV 22
Db 582 LKQQQLGIWCGSGKLCITTA 603

RESULT 7
ENV_HV1BN
ID ENV_HV1BN STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21098; AAA44221.1; -
CC PIR; A31667; VCLJBR.
CC PDB; 1IM7; 23-OCT-02.
CC HIV; M21098; ENV$SRVA.
CC InterPro; IPR000328; Env_GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure. 1 30
FT SIGNAL 31 507 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.
```

```
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 155 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 339 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 439 BY SIMILARITY.
FT DISULFID 383 412 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 60.7%; Score 74; DB 1; Length 852;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNOQRLLNSWGCKGRIVCVTSV 22
Db 583 LKQQQLGIWCGSGKLCITTA 604

RESULT 8
ENV_HV1S3
ID ENV_HV1S3 STANDARD; PRT; 852 AA.
AC F19549;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RT J. Virol. 64:4016-4020(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AY352275; RAQ17031.1; -.  
DR PDB; 1MEQ; 11-DEC-02.  
DR HIV; M38427; ENVSSF33.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
3D-structure.

FT SIGNAL 1 31 BY SIMILARITY.  
FT CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 507 852 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 206 BY SIMILARITY.  
FT DISULFID 125 197 BY SIMILARITY.  
FT DISULFID 130 156 BY SIMILARITY.  
FT DISULFID 219 248 BY SIMILARITY.  
FT DISULFID 229 240 BY SIMILARITY.  
FT DISULFID 297 331 BY SIMILARITY.  
FT DISULFID 377 439 BY SIMILARITY.  
FT DISULFID 384 412 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 852 AA; 96663 MW; EE7BBF8D23C9910D CRC64;

Query Match 60.7%; Score 74; DB 1; Length 852;  
Best Local Similarity 54.5%; Pred. No. 0.00054;  
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQLLSWGCKGLVCVTSV 22  
Db 583 LRDQQLGIWGCGKLCITTV 604

RESULT 9

ENV\_HVIMF STANDARD; PRT; 853 AA.  
AC F19551;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV  
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11704;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90317877; PubMed=1695254;  
RA Stevenson M., Haggerty S., Lamonica C., Mann A.M., Meier C.,  
RA Wasiaak A.;  
RT "Cloning and characterization of human immunodeficiency virus type 1  
RT variants diminished in the ability to induce syncytium-independent  
RT cytolysis";  
RL J. Virol. 64:3792-3803 (1990).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M33943; AAA44850.1; -.  
DR PDB; 1A1K; 16-JUN-97.  
DR HIV; M33943; ENV\$MFA.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
3D-structure. 1 30  
FT SIGNAL 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 203 BY SIMILARITY.  
FT DISULFID 126 194 BY SIMILARITY.  
FT DISULFID 131 157 BY SIMILARITY.  
FT DISULFID 216 245 BY SIMILARITY.  
FT DISULFID 226 237 BY SIMILARITY.  
FT DISULFID 294 329 BY SIMILARITY.  
FT DISULFID 376 443 BY SIMILARITY.  
FT DISULFID 383 416 BY SIMILARITY.  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 672 672 N-LINKED (GLCNAC. .) (POTENTIAL).



FT	CARBOHYD	296	296	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	331	331	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	353	353	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	384	384	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	390	390	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	402	402	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	441	441	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	445	445	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	458	458	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	459	459	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	462	462	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	608	608	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	613	613	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	622	622	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	634	634	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	671	671	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	SEQUENCE	853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;			

Query Match 60.7%; Score 74; DB 1; Length 853;  
 Best Local Similarity 54.5%; Pred. No. 0.00054;  
 Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY	1	LNQORLLNSGCKGRIVCYTSV 22
DB	584	LKDQQLGIWGCXGLICTTV 605
		:           :   :   :

RESULT 11  
 ENV\_HV1A2 STANDARD; PRT; 855 AA.

ID	ENV_HV1A2	AC	P03378;
DT	21-JUL-1986	(Rel. 01, Created)	
DT	21-JUL-1986	(Rel. 01, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane		
DE	glycoprotein (GP120); Transmembrane glycoprotein (GP41)]		
OS	Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11685;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85090453; PubMed=2578227;		
RA	Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,		
RA	Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,		
RA	Levy J.A., Dina D., Luciw P.A.;		
RT	"Nucleotide sequence and expression of an AIDS-associated retrovirus		
RT	(ARV-2).";		
RL	Science 227:484-492(1985).		

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR	EMBL; K02007; AAB59882.1; -
DR	PIR; A03976; VCLJAJ2.
DR	HIV; K02007; ENV\$SF2.
DR	InterPro; IPR00328; Env GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
DR	Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW	Signal.
FT	SIGNAL
FT	CHAIN
FT	CHAIN
FT	DISULFID
FT	DISULFID

1 29  
 30 509  
 510 855  
 53 73  
 118 208  
 BY SIMILARITY.

EXTERIOR MEMBRANE GLYCOPROTEIN.  
 TRANSMEMBRANE GLYCOPROTEIN.  
 BY SIMILARITY.

```

FT DISULFID 125 199 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 221 250 BY SIMILARITY.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 239 333 BY SIMILARITY.
FT DISULFID 380 442 BY SIMILARITY.
FT DISULFID 387 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 60.7%; Score 74; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQLLNSWGCKGLVCYTSV 22
Db 586 LKQQLLGWCGSKGLICTTAV 607

RESULT 12
ENV_HV10Y
ID ENV_HV10Y STANDARD; PRT; 855 AA.
AC P20888;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11699;
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989)
CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M26727; AAA83397.1; -.
CC HIV; M26727; ENVSOVI.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Glycoprotein; Transmembrane;
CC Signal.
CC SIGNAL. 1 29
CC CHAIN 30 509
CC CHAIN 510 855
CC DISULFID 53 73 BY SIMILARITY.
CC DISULFID 118 210 BY SIMILARITY.
CC DISULFID 125 201 BY SIMILARITY.
CC DISULFID 130 162 BY SIMILARITY.
CC DISULFID 223 252 BY SIMILARITY.
CC DISULFID 233 244 BY SIMILARITY.
CC DISULFID 301 335 BY SIMILARITY.
CC DISULFID 381 442 BY SIMILARITY.
CC DISULFID 388 415 BY SIMILARITY.
CC CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 60.7%; Score 74; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LNQQLLNSWGCKGLVCYTSV 22
Db 586 LKQQLLGWCGSKGLICTTAV 607

RESULT 13
ENV_HV126
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
-----
```

DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
GN ENV.  
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87248097; PubMed=3036660;  
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,  
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,  
RA Sanchez-Pescador R.;  
RT "Molecular characterization of human immunodeficiency virus from  
RT Zaire: nucleotide sequence analysis identifies conserved and variable  
RT domains in the envelope gene.";  
RL Gene 52:71-82(1987).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: K03458; AAA45380.1; --  
DR PIR; D26192; VCLJZR.  
DR HIV; K03458; ENV\$Z6.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW Signal.  
FT CHAIN 1 30  
FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 207 BY SIMILARITY.  
FT DISULFID 125 198 BY SIMILARITY.  
FT DISULFID 130 259 BY SIMILARITY.  
FT DISULFID 220 249 BY SIMILARITY.  
FT DISULFID 230 241 BY SIMILARITY.  
FT DISULFID 298 332 BY SIMILARITY.  
FT DISULFID 378 444 BY SIMILARITY.  
FT DISULFID 385 417 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;  
Query Match 60.7%; Score 74; DB 1; Length 855;  
Best Local Similarity 54.5%; Pred.No. 0.00054;  
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
QY 1 LNOQRLLNGWGCKGRVLCVTSV 22  
DB 586 LKQQQLLGWCSGKLCITTV 607  
RESULT 14  
ENV HV1B1 STANDARD; PRT; 856 AA.  
AC P03375; 1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11678;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
RA Lautenberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C.,  
RA Wong-Staal P.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
RN [2]  
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=90285159; PubMed=2355006;  
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,  
RA Gregory T.J.;  
RT "Assignment of intrachain disulfide bonds and characterization of  
RT potential glycosylation sites of the type 1 recombinant human  
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in  
RT Chinese hamster ovary cells.";  
RL J. Biol. Chem. 265:10373-10382(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M15654; AAA44205.1; --  
DR PIR; A03973; VCLJH3.  
DR HIV; M15654; ENV\$BH102.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW Signal.  
FT CHAIN 1 30  
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74  
FT DISULFID 119 205  
FT DISULFID 126 196  
FT DISULFID 131 157



```

Best Local Similarity 54.5%; Pred. No. 0.00054;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LNQQRLNMGCKGRINCYTSV 22
    | : | | | | : | : | : |
Db 587 LKQQLLGIGWCSGKLICTTAV 608

```

Search completed: May 7, 2004, 18:32:32  
Job time : 8.21393 secs

CRASH INVESTIGATION

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:22:06 ; Search time 32.8358 Seconds  
(without alignments)  
211.397 Million cell updates/sec

Title: US-09-147-362A-14  
Perfect score: 122  
Sequence: 1 LNQQRLLNSWGCKRLVCYTSV 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	86.1	116	15 Q7ZJN9	Q7zjn9 human immun
2	105	86.1	216	15 Q9IEC5	Q9iec5 human immun
3	105	86.1	219	15 Q9IEB6	Q9ieb6 human immun
4	105	86.1	890	15 Q8Q7G2	Q8q7g2 human immun
5	103	84.4	219	15 Q9IEC8	Q9iec8 human immun
6	102	83.6	130	15 Q9IHU9	Q9ihu9 human immun
7	102	83.6	872	15 Q8Q7H0	Q8q7h0 human immun
8	102	83.6	882	15 Q8Q7F9	Q8q7f9 human immun
9	102	83.6	887	15 Q8Q7H6	Q8q7h6 human immun
10	102	83.6	887	15 Q8Q7G9	Q8q7g9 human immun
11	101	82.8	135	15 Q9DQL9	Q9dql9 human immun
12	101	82.8	242	15 Q9IE31	Q9ie31 human immun
13	99	81.1	115	15 Q7ZJN8	Q7zjn8 human immun
14	99	81.1	116	15 Q40459	Q40459 human immun
15	99	81.1	118	15 Q40451	Q40451 human immun
16	99	81.1	120	15 Q9IHU2	Q9ihu2 human immun

17	99	81.1	131	15 Q9WR05	Q9wr05 human immun
18	99	81.1	131	15 Q9IHU0	Q9ihu0 human immun
19	99	81.1	137	15 Q9IHV5	Q9ihv5 human immun
20	99	81.1	146	15 Q9WRV2	Q9wrv2 human immun
21	99	81.1	213	15 Q9IEC3	Q9iec3 human immun
22	99	81.1	214	15 Q9DIK3	Q9dik3 human immun
23	99	81.1	214	15 Q9IE96	Q9ie96 human immun
24	99	81.1	216	15 Q9IEA5	Q9iea5 human immun
25	99	81.1	342	15 Q9IEE9	Q9iee9 human immun
26	99	81.1	532	15 Q9IEE9	Q9iee9 human immun
27	99	81.1	544	15 Q9IED9	Q9ied9 human immun
28	99	81.1	548	15 Q9IED6	Q9ied6 human immun
29	99	81.1	867	15 Q8Q7G8	Q8q7g8 human immun
30	99	81.1	871	15 Q57073	Q57073 human immun
31	99	81.1	871	15 Q57074	Q57074 human immun
32	99	81.1	871	15 Q8Q7I2	Q8q7i2 human immun
33	99	81.1	872	15 Q900Y5	Q900y5 human immun
34	99	81.1	872	15 Q57072	Q57072 human immun
35	99	81.1	900	15 Q9QN28	Q9qnz8 human immun
36	98	80.3	114	15 Q40448	Q40448 human immun
37	98	80.3	116	15 Q11941	Q11941 human immun
38	98	80.3	116	15 Q40458	Q40458 human immun
39	98	80.3	116	15 Q40449	Q40449 human immun
40	98	80.3	117	15 Q9Q6F2	Q9q6f2 human immun
41	98	80.3	124	15 Q9IHU7	Q9ihu7 human immun
42	98	80.3	132	15 Q9IHV0	Q9ihv0 human immun
43	98	80.3	134	15 Q9IHV4	Q9ihv4 human immun
44	98	80.3	155	15 Q8J3Q4	Q8j3q4 human immun
45	98	80.3	158	15 Q8J3N6	Q8j3n6 human immun

ALIGNMENTS

RESULT 1

Q7ZJN9 PRELIMINARY; PRT; 116 AA.  
AC Q7ZJN9;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]\_TaxID=11676;  
RP SEQUENCE FROM N.A.  
RC STRAIN=08692A;  
RA Swanson P.A.; Devare S.G.; Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
(Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
integrase, and env gp41.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214120; AAO61840.1; -;  
DR GO; GO:0019031; C:Viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 116 116  
SQ SEQUENCE 116 AA; 13865 MW; AFIDFPCS7E059061 CRC64;

Query Match 86.1%; Score 105; DB 15; Length 116;  
Best Local Similarity 81.8%; Pred. No. 1.8e-09;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLLNSWGCKRLVCYTSV 22

Db 27 LNQQRLLNSWGCKRLVCYTSV 48

RESULT 2

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5

```
AC Q9IEC5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF07;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ236394; CAB96243.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match      86.1%; Score 105; DB 15; Length 216;
Best Local Similarity 81.8%; Pred. No. 3.4e-09;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db :|:|||||
42 IQNQQLLNSWGCKGRLVCYTSV 63

RESULT 3
Q9IEB6
ID Q9IEB6 PRELIMINARY; PRT; 219 AA.
AC Q9IEB6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF14;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ236404; CAB96252.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match      86.1%; Score 105; DB 15; Length 219;
Best Local Similarity 81.8%; Pred. No. 3.4e-09;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db :|:|||||
52 IQNQQLLNSWGCKGRLVCYTSV 73
```

```
RESULT 4
Q9Q7G2
ID Q9Q7G2 PRELIMINARY; PRT; 890 AA.
AC Q9Q7G2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97US08692A;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282 (2002).
DR EMBL; AF383259; AAL98881.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 890 AA; 100610 MW; BBE816BC5C45EE23 CRC64;

Query Match      86.1%; Score 105; DB 15; Length 890;
Best Local Similarity 81.8%; Pred. No. 1.4e-08;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
Db :|:|||||
615 IQNQQLLNSWGCKGRLVCYTSV 636

RESULT 5
Q9IEC8
ID Q9IEC8 PRELIMINARY; PRT; 219 AA.
AC Q9IEC8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF02;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ236391; CAB96240.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;
```



```
Query Match      84.4%; Score 103; DB 15; Length 219;
Best Local Similarity 77.3%; Pred. No. 7.3e-09;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 51 IQNQQLNSWGCKGRIVCYTSV 72

RESULT 6
Q9IHU9 PRELIMINARY; PRT; 130 AA.
AC Q9IHU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229235; AAF71912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; F:structural molecule activity; IEA.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 130 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match      83.6%; Score 102; DB 15; Length 130;
Best Local Similarity 77.3%; Pred. No. 6.4e-09;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 35 IQNQQLNSWGCKGRIVCYTSV 56

RESULT 7
Q8Q7H0 PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383251; AAL98873.1; -.

Query Match      83.6%; Score 102; DB 15; Length 882;
Best Local Similarity 77.3%; Pred. No. 4.2e-08;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 608 IQNQQLNSWGCKGRIVCYTAV 629

RESULT 9
Q8Q7H6 PRELIMINARY; PRT; 887 AA.
AC Q8Q7H6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA100;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383262; AAL98884.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 6278CA5D0C09C6D4 CRC64;

Query Match      83.6%; Score 102; DB 15; Length 882;
Best Local Similarity 77.3%; Pred. No. 4.2e-08;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNQQRLNSWGCKGRLVCYTSV 22
DB 608 IQNQQLNSWGCKGRIVCYTAV 629

RESULT 9
Q8Q7H6 PRELIMINARY; PRT; 887 AA.
AC Q8Q7H6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA100;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383262; AAL98884.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 6278CA5D0C09C6D4 CRC64;
```





THIS PAGE ON ANY OTHER

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:51:11 ; Search time 59.7612 Seconds  
(without alignments)  
132.382 Million cell updates/sec

Title: US-09-147-362A-15

Perfect score: 149  
Sequence: 1 RALETLLNQRLNLSWGCKRLVCYTSV 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	149	100.0	28	2 AAW80473	Aaw80473 Peptide d
2	127	85.2	33	3 AAB12231	Aab12231 Partial s
3	127	85.2	33	3 AAB12212	Aab12212 Partial s
4	127	85.2	40	2 AAW07346	Aaw07346 Partial s
5	127	85.2	40	2 AAW07352	Aaw07352 Partial s
6	126	84.6	32	2 AAW80469	Aaw80469 Peptide d
7	125	83.9	33	3 AAB12236	Aab12236 Partial s
8	125	83.9	40	2 AAW07343	Aaw07343 Partial s
9	122	81.9	22	2 AAW80472	Aaw80472 Peptide d
10	121	81.2	33	3 AAB12210	Aab12210 Partial s
11	121	81.2	33	3 AAB12221	Aab12221 Partial s
12	121	81.2	33	3 AAB12213	Aab12213 Partial s
13	121	81.2	41	2 AAW07353	Aaw07353 Partial s
14	121	81.2	42	2 AAW07350	Aaw07350 Partial s
15	121	81.2	116	2 AAY05555	Aay05555 HIV-1 gro
16	121	81.2	117	2 AAY05548	Aay05548 HIV-1 gro
17	121	81.2	715	2 AAY05625	Aay05625 HIV-1 gro
18	120	80.5	32	2 AAW80470	Aaw80470 Peptide d
19	120	80.5	33	3 AAB12211	Aab12211 Partial s
20	120	80.5	33	3 AAB12222	Aab12222 Partial s
21	120	80.5	33	3 AAB12214	Aab12214 Partial s
22	120	80.5	41	2 AAW07351	Aaw07351 Partial s
23	120	80.5	113	2 AAY05559	Aay05559 HIV-1 gro
24	120	80.5	115	2 AAY05557	Aay05557 HIV-1 gro
25	120	80.5	200	3 AAY77373	Aay77373 HIV-1 gro

26	120	80.5	215	2 AAY09499	Aay09499 HIV-1 Gro
27	120	80.5	215	2 AAY06983	Aay06983 Recombina
28	120	80.5	215	3 AAY77374	Aay77374 HIV-1 Gro
29	120	80.5	245	2 AAY09493	Aay09493 HIV-1 Gro
30	120	80.5	245	2 AAY06977	Aay06977 Recombina
31	120	80.5	245	3 AAY77369	Aay77369 HIV-1 gro
32	120	80.5	281	2 AAY09507	Aay09507 HIV-1 Gro
33	120	80.5	373	2 AAY09495	Aay09495 HIV-1 Gro
34	120	80.5	373	2 AAY06979	Aay06979 Recombina
35	120	80.5	460	2 AAY09500	Aay09500 HIV-1 Gro
36	120	80.5	460	2 AAY06984	Aay06984 Recombina
37	120	80.5	460	3 AAY77375	Aay77375 HIV-1 gro
38	120	80.5	474	3 AAY77371	Aay77371 HIV-1 Gro
39	120	80.5	488	2 AAY09504	Aay09504 HIV-1 Gro
40	120	80.5	490	2 AAY09494	Aay09494 HIV-1 Gro
41	120	80.5	490	2 AAY06978	Aay06978 Recombina
42	120	80.5	490	3 AAY77370	Aay77370 HIV-1 gro
43	120	80.5	526	2 AAY09505	Aay09505 HIV-1 Gro
44	120	80.5	618	2 AAY09496	Aay09496 HIV-1 Gro
45	120	80.5	618	2 AAY06980	Aay06980 Recombina

ALIGNMENTS

RESULT 1  
AAW80473  
ID AAW80473 standard; peptide; 28 AA.

XX AC AAW80473;  
XX XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX XX  
KW Group O human immune deficiency virus; HIV; detection; infection.

OS Synthetic.  
OS Human immunodeficiency virus 1.

XX WO9845323-A1.

XX PD 15-OCT-1998.

XX PF 06-APR-1998; 98WO-FR000691.

XX PR 09-APR-1997; 97FR-00004356.

XX PR 24-FEB-1998; 98FR-00002212.

XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;

XX XX WPI; 1998-583190/49.

XX PT New synthetic peptide(s) - useful for, e.g. detecting infection by human immune deficiency virus of group O.

XX PS Claim 6; Page 44; 55pp; French.

XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences connected around short highly conserved sequences present in isolates of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)

XX CC (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 28 AA;

Query Match 100.0%; Score 149; DB 2; Length 28;

```
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RALETLLNQORLLNSWGCKGRVLCYTSV 28
    |||||:|||||:|||||:|||||:|||||
Db 1 RALETLLNQORLLNSWGCKGRVLCYTSV 28
    |||||:|||||:|||||:|||||:|||||

RESULT 2
AAB12231
ID AAB12231 standard; peptide; 33 AA.
XX AC AAB12231;
XX DT 12-SEP-2003 (revised)
XX DT 10-NOV-2000 (first entry)
XX DE Partial sequence of HIV-1 strain MAN gp41 immunodominant region.
XX KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
XX KW acquired immunodeficiency syndrome; group O HIV; gp41; MAN.
XX OS Human immunodeficiency virus 1.
XX PN EP1013766-A2.
XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX PF New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
XX PS Example 1; Fig 1; 52pp; English.
XX CC The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain MAN. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. MAN is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment, which
CC in turn was used to derive a consensus sequence peptide; peptide 147
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 33 AA;

Query Match 85.2%; Score 127; DB 3; Length 33;
Best Local Similarity 85.2%; Pred. No. 2e-10;
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLLNQORLLNSWGCKGRVLCYTSV 28
    |||||:|||||:|||||:|||||:|||||
Db 4 ALETLLNQORLLNSWGCKGRVLCYTSV 30
    |||||:|||||:|||||:|||||:|||||

RESULT 4
AAW07346
ID AAW07346 standard; peptide; 40 AA.
XX AC AAW07346;
XX DT 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 03-JUN-1997 (first entry)
XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).
XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
XX KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
XX KW primer; hybridisation; amplification; PCR; polymerase chain reaction;
XX KW immunogen; antibody.
XX OS Human immunodeficiency virus 1.
XX PN W09627013-A1.
```

```

XX 27-FEB-1995;    95FR-00002336.
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX Simon F, Saragosti S, Loussetajaka I, Ly T, Chauxbaudier M;
XX WPI; 1996-412779/41.
XX
XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and
XX antibodies - useful for diagnosis, screening and typing, or as
XX immunogens.
XX
XX Claim 12; Page 46; 7lpp; French.
XX
XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
XX into 2 major groups based on the nucleotide sequences of the envelop gene
XX (env): group M containing sub-groups A-G, and group O containing the
XX strains AN770 and MVP180. The invention relates to the discovery of
XX several new strains of HIV-1 which can be placed in group O, based on the
XX partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
XX and AA07329-64). The novel strains have been deposited as retroviruses
XX CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
XX (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
XX the strain BCF13 and corresponds to a fragment of the gp41 protein
XX encoded by the env gene. The nucleic acids can be used to detect gp. O
XX HIV-1 strains by hybridisation or (as primers) by gene amplification,
XX also for screening and typing of such strains. Peptides encoded by the
XX nucleic acids can be used as immunogens to raise Ab for detecting gp. O
XX HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
XX 2003 to standardise OS field)
XX SQ Sequencé 40 AA;
Query Match      85.2%; Score 127; DB 2; Length 40;
Best Local Similarity 85.2%; Pred. No. 2.5e-10;
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 ALSTLLNQQLLSWGCKGLVCYTSSV 28
DB 9 ALSTLIQNQQLLSWGCKGLVCYTSSV 35
|||||:|||||||
ALSTLIQNQQLLSWGCKGLVCYTSSV 35
RESULT 6
ID AAW80469 standard; peptide; 32 AA.
XX AC AAW80469;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX
XX Peptide derived from a conserved sequence of group O human HIV.
XX DE Group O human immune deficiency virus; HIV; detection; infection.
XX KW Synthetic.
XX OS Human immunodeficiency virus 1.
XX OS WO9845323-A1.
XX PN
XX PD 15-OCT-1998.
XX PF 06-APR-1998; 98WO-FR000691.
XX PR 09-APR-1997; 97FR-00004356.
XX PR 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX PA
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;

```

```

XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
XX Claim 6; Page 44; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immunodeficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 32 AA;
SQ
    Query Match      84.6%; Score 126; DB 2; Length 32;
    Best Local Similarity 85.2%; Pred. No. 2.7e-10;
    Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
    QY 2 ALETLNQRLNSWGCKGRVLCVTSV 28
    Db 1 ALETLNQRLNSWGCKGRVLCVTSV 27
    RESULT 7
    AAB12236
    ID AAB12236 standard; peptide; 33 AA.
    AC AAB12236;
    XX
    XX 12-SEP-2003 (revised)
    DT 10-NOV-2000 (first entry)
    XX
    XX Partial sequence of HIV-1 strain ESS gp41 immunodominant region.
    XX
    XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
    KW acquired immunodeficiency syndrome; group O HIV; gp41; ESS.
    XX
    XX Human immunodeficiency virus 1.
    XX
    XX EP1013766-A2.
    XX
    XX 28-JUN-2000.
    XX
    XX 29-NOV-1999; 99BP-00309491.
    XX
    XX 30-NOV-1998; 98US-0110292P.
    PR 08-FEB-1999; 99US-0119138P.
    PR 04-NOV-1999; 99US-00433428.
    XX
    XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
    PA
    XX
    XX De Leys R, Zheng J;
    PI
    XX WPI; 2000-402205/35.
    DR
    XX
    XX New antigenic peptides and peptide functional derivatives, useful for
    PT detection of antibodies produced in response to human immunodeficiency
    PT virus group O antibodies.
    PT
    XX
    XX Example 1; Fig 1; 52pp; English.
    PS
    XX
    XX The present sequence is a partial gp41 protein of Human Immunodeficiency
    CC Virus Type 1 (HIV-1) strain ESS. HIV is the principle aetiological agent
    CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope
    CC protein, and may be used as an antigen for the detection of antibodies
    CC produced in response to HIV infection. ESS is a member of HIV group O
    CC (outlier). The present sequence is the immunodominant region of gp41.
    CC This sequence was used in a sequence homology alignment, which in turn
    CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).
    CC

CC (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 33 AA;
SQ
    Query Match      83.9%; Score 125; DB 3; Length 33;
    Best Local Similarity 81.5%; Pred. No. 3.9e-10;
    Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
    QY 2 ALETLNQRLNSWGCKGRVLCVTSV 28
    Db 4 ALETLNQRLNSWGCKGRVLCVTSV 30
    RESULT 8
    AAW07343
    ID AAW07343 standard; peptide; 40 AA.
    XX
    AC AAW07343;
    XX
    XX 16-OCT-2003 (revised)
    DT 25-MAR-2003 (revised)
    DT 03-JUN-1997 (first entry)
    XX
    XX Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (ESS) .
    XX
    XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
    KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
    KW primer; hybridisation; amplification; PCR; polymerase chain reaction;
    KW immunogen; antibody.
    XX
    XX Human immunodeficiency virus 1.
    XX
    XX WO9627013-A1.
    PN
    XX 06-SEP-1996.
    PD
    XX 26-FEB-1996; 96WO-FR000294.
    PF
    XX 27-FEB-1995; 95FR-00002236.
    PR
    XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
    PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
    XX
    XX Simon F, Saragosti S, Loussertajaka I, Ly T, Chaixbaudier M;
    PI
    XX WPI; 1996-412779/41.
    DR N-PSDB; AAT44918.
    DR
    XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and
    PT antibodies - useful for diagnosis, screening and typing, or as
    PT immunogens.
    PT
    XX
    XX Claim 12; Page 33; 71pp; French.
    PS
    XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
    CC into 2 major groups based on the nucleotide sequences of the envelop gene
    CC (env): group M containing sub-groups A-G, and group O containing the
    CC strains AMT70 and WVP5180. The invention relates to the discovery of
    CC several new strains of HIV-1 which can be placed in group O, based on the
    CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
    CC and AAW07329-64). The novel strains have been deposited as retroviruses
    CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
    CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
    CC the strain BCF02 (ESS) and corresponds to a fragment of the gp41 protein
    CC encoded by the env gene. The nucleic acids can be used to detect gp. O
    CC HIV-1 strains by hybridisation or (as primers) by gene amplification,
    CC also for screening and typing of such strains. Peptides encoded by the
    CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O
    CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
    CC 2003 to standardise OS field)
    XX
    XX Sequence 40 AA;
    SQ

```



```
Query Match      83.9%; Score 125; DB 2; Length 40;
Best Local Similarity 81.5%; Pred. No. 4.7e-10;
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28
    |||||: |:|||:|||||:|||||
Db 9 ALETLNQORLLNSWGCKGRLVCYTSV 35

RESULT 9
AAW80472
ID AAW80472 standard; peptide; 22 AA.
XX
AC AAW80472;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
XX Peptide derived from a conserved sequence of group O human HIV.
DE
XX Group O human immune deficiency virus; HIV; detection; infection.
KW
XX Synthetic.
OS
OS Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
PN
XX
XX 15-OCT-1998.
PD
XX
XX 06-APR-1998; 98WO-FR000691.
PF
XX
XX 09-APR-1997; 97FR-00004356.
PR
XX 24-FEB-1998; 98FR-00002212.
PR
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
PA
XX
XX Chenebaux DMB, Delagneau JH, Gabelle SJX, Rieunier FY;
PI
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
PT
XX
XX Claim 6; Page 44; 55pp; French.
PS
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
CC
XX
XX Sequence 22 AA;
SQ

Query Match      81.9%; Score 122; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LNQORLLNSWGCKGRLVCYTSV 28
    |||||:|||||:|||||:|||||
Db 1 LNQORLLNSWGCKGRLVCYTSV 22

RESULT 10
AAW12210
ID AAW12210 standard; peptide; 33 AA.
XX
XX AAW12210;
AC
XX
XX 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
PF
XX

Partial sequence of HIV-1 strain BCF09 gp41 immunodominant region.
HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
acquired immunodeficiency syndrome; group O HIV; gp41; BCF09.
Human immunodeficiency virus 1.
EP1013766-A2.
28-JUN-2000.
29-NOV-1999; 99EP-00309491.
30-NOV-1998; 98US-0110292P.
08-FEB-1999; 99US-0119138P.
04-NOV-1999; 99US-00433428.
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
De Leys R, Zheng J;
WPI; 2000-402205/35.
New antigenic peptides and peptide functional derivatives, useful for
detection of antibodies produced in response to human immunodeficiency
virus group O antibodies.
Example 1; Fig 1; 52pp; English.
The present sequence is a partial gp41 protein of Human Immunodeficiency
Virus Type 1 (HIV-1) strain BCF09. HIV is the principle aetiological
agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
envelope protein, and may be used as an antigen for the detection of
antibodies produced in response to HIV infection. BCF09 is a member of
HIV group O (outlier). The present sequence is the immunodominant region
of gp41. This sequence was used in a sequence homology alignment, which
in turn was used to derive a consensus sequence peptide: peptide 147
(AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 33 AA;
SQ

Query Match      81.2%; Score 121; DB 3; Length 33;
Best Local Similarity 81.5%; Pred. No. 1.4e-09;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28
    |||||: |:|||:|||||:|||||
Db 4 ALETLNQORLLNSWGCKGRLVCYTSV 30

RESULT 11
AAB12221
ID AAB12221 standard; peptide; 33 AA.
XX
XX AAB12221;
AC
XX
XX 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
DT
XX
XX Partial sequence of HIV-1 strain ABT2156 gp41 immunodominant region.
HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
acquired immunodeficiency syndrome; group O HIV; gp41; ABT2156.
Human immunodeficiency virus 1.
EP1013766-A2.
28-JUN-2000.
29-NOV-1999; 99EP-00309491.
XX
```

```
PR 30-NOV-1998; 98US-0110292P.
PR 08-FEB-1999; 99US-0119138P.
PR 04-NOV-1999; 99US-00433428.
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI De Leys R, Zheng J;
XX
DR WPI; 2000-402205/35.
XX
PT New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
XX
PS Example 1; Fig 1; 52pp; English.
XX
CC The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain ABT2156. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. ABT2156 is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment, which
CC in turn was used to derive a consensus sequence peptide: peptide 147
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 33 AA;

Query Match 81.2%; Score 121; DB 3; Length 33;
Best Local Similarity 81.5%; Pred. No. 1.4e-09;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQRLNSWGCKGRVLCYTSV 28
   |||||:|:|||||
Db 4 ALETLLNQRLNSWGCKGRVLCYTSV 30

RESULT 13
AAW07353
ID AAW07353 standard; peptide; 41 AA.
XX
AC AAW07353;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 03-JUN-1997 (first entry)
XX
DE Partial sequence of gp41 from HIV-1 gp. O strain BCF14.
XX
KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
KW primer; hybridisation; amplification; PCR; polymerase chain reaction;
KW immunogen; antibody.
XX
OS Human immunodeficiency virus 1.
XX
PN WO9627013-A1.
XX
PD 06-SEP-1996.
XX
PF 26-FEB-1996; 96WO-FR000294.
XX
PR 27-FEB-1995; 95FR-00002236.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX
PI Simon F, Saragosti S, Loussertajaka I, Ly T, Chaixbaudier M;
XX WPI; 1996-412779/41.
XX
DR New strains of HIV-1 group O, related DNA fragments, peptide(s) and
PT antibodies - useful for diagnosis, screening and typing, or as
PT immunogens.
XX
PS Claim 12; Page 46; 71pp; French.
XX
CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
CC into 2 major groups based on the nucleotide sequences of the envelop gene
CC (env): group M containing sub-groups A-G, and group O containing the
CC strains ANT70 and MVP5180. The invention relates to the discovery of
CC several new strains of HIV-1 which can be placed in group O, based on the
CC partial sequences of the C2V3-env. gp41 and gag genes (see AAT4907-39
CC and AAW07329-64). The novel strains have been deposited as retroviruses
CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
CC the strain BCF14 and corresponds to a fragment of the gp41 protein
```

```
PR 30-NOV-1998; 98US-0110292P.
PR 08-FEB-1999; 99US-0119138P.
PR 04-NOV-1999; 99US-00433428.
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI De Leys R, Zheng J;
XX
DR WPI; 2000-402205/35.
XX
PT New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
XX
PS Example 1; Fig 1; 52pp; English.
XX
CC The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain ABT2156. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. ABT2156 is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment, which
CC in turn was used to derive a consensus sequence peptide: peptide 147
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 33 AA;

Query Match 81.2%; Score 121; DB 3; Length 33;
Best Local Similarity 81.5%; Pred. No. 1.4e-09;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQRLNSWGCKGRVLCYTSV 28
   |||||:|:|||||
Db 4 ALETLLNQRLNSWGCKGRVLCYTSV 30

RESULT 12
AAB12213
ID AAB12213 standard; peptide; 33 AA.
XX
AC AAB12213;
XX
DT 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX
DE Partial sequence of HIV-1 strain BCF14 gp41 immunodominant region.
XX
KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF14.
XX
OS Human immunodeficiency virus 1.
XX
PN EP1013766-A2.
XX
PD 28-JUN-2000.
XX
PF 29-NOV-1999; 99EP-00309491.
XX
PR 30-NOV-1998; 98US-0110292P.
PR 08-FEB-1999; 99US-0119138P.
PR 04-NOV-1999; 99US-00433428.
XX
PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI De Leys R, Zheng J;
XX
DR WPI; 2000-402205/35.
XX
PT New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
XX
```

CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
CC also for screening and typing of such strains. Peptides encoded by the  
CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
CC 2003 to standardise OS field)  
XX  
SQ Sequence 41 AA;

Query Match 81.2%; Score 121; DB 2; Length 41;  
Best Local Similarity 81.5%; Pred. No. 1.7e-09;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQQLLNLSWGCKGRVLCVTSV 28  
|||||:|:|||||||  
Db 9 ALETLLNQQLLNLSWGCKGRVLCVTSV 35

RESULT 14  
AAW07350  
ID AAW07350 standard; peptide; 42 AA.  
XX  
AC AAW07350;

DT 16-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 03-JUN-1997 (first entry)

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF09.

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
KW immunogen; antibody.

XX Human immunodeficiency virus 1.

XX WO9627013-A1.

XX 06-SEP-1996.

XX 26-FEB-1996; 96WO-FR000294.

XX 27-FEB-1995; 95FR-00002236.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
PT antibodies - useful for diagnosis, screening and typing, or as  
PT immunogens.

XX Claim 12; Page 45; 71pp; French.

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
CC into 2 major groups based on the nucleotide sequences of the envelop gene  
CC (env): group M containing sub-groups A-G, and group O containing the  
CC strains ANT70 and MVP5180. The invention relates to the discovery of  
CC several new strains of HIV-1 which can be placed in group O, based on the  
CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
CC and AAW07329-64). The novel strains have been deposited as retroviruses  
CC CMCN I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (WAN)), 1547  
CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
CC the strain BCF09 and corresponds to a fragment of the gp41 protein  
CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
CC also for screening and typing of such strains. Peptides encoded by the  
CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-

CC 2003 to standardise OS field)  
XX  
SQ Sequence 42 AA;

Query Match 81.2%; Score 121; DB 2; Length 42;  
Best Local Similarity 81.5%; Pred. No. 1.8e-09;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQQLLNLSWGCKGRVLCVTSV 28  
|||||:|:|||||||  
Db 9 ALETLLNQQLLNLSWGCKGRVLCVTSV 35

RESULT 15  
AAV05555  
ID AAV05555 standard; protein; 116 AA.  
XX  
AC AAV05555;

DT 17-OCT-2003 (revised)  
DT 19-JUL-1999 (first entry)

XX HIV-1 group O isolate 320 gp41 antigen.

XX HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine;  
KW diagnosis; AIDS.

XX Human immunodeficiency virus 1.

XX WO9904011-A2.

XX 28-JAN-1999.

XX 20-JUL-1998; 98WO-EP004522.

XX 18-JUL-1997; 97EP-00870110.

XX (INNO-) INNOGENETICS NV.

XX Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;

XX WPI; 1999-132255/11.

XX N-PSDB; AAX25163.

XX New isolated HIV-1 group O strains - used to produce polynucleotides,  
PT antigens and antibodies for use in diagnosis and in vaccines for  
PT prevention of HIV-1 infection.

XX Claim 3; Fig 6; 162pp; English.

XX The present sequence is an antigen of the gp41 protein of HIV-1 group O  
CC (Outlier) strain 320, an isolate from Tchaad. The invention relates to  
CC new HIV-1 group O antigens (see AAY05546-625), and the use of these  
CC antigens, or nucleic acids encoding them (see AAX25154-80), in the  
CC diagnosis and prophylaxis of AIDS. They can be used as reagents for  
CC detecting HIV-1 group O infection and for differentiating different types  
CC of HIV-1 group O infection. Vaccines that provide protective immunity  
CC against HIV-1 infection, in particular against HIV-1 group O infection,  
CC comprise at least one HIV-1 type O antigen, a nucleic acid encoding such  
CC an antigen, a virus-like particle comprising such an antigen, or an  
CC attenuated form of an HIV-1 type O strain. The invention also relates to  
CC new HIV-1 group O strains, mostly from patients from Cameroon and its  
CC neighbouring countries. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 116 AA;

Query Match 81.2%; Score 121; DB 2; Length 116;  
Best Local Similarity 81.5%; Pred. No. 5.1e-09;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQQLLNLSWGCKGRVLCVTSV 28  
|||||:|:|||||||  
Db 29 ALETLLNQQLLNLSWGCKGRVLCVTSV 55

Search completed: May 7, 2004, 18:31:55  
Job time : 59.7612 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:29:27 ; Search time 17.5522 Seconds  
(without alignments)  
82.356 Million cell updates/sec

Title: US-09-147-362A-15  
Perfect score: 149  
Sequence: 1 RALETLLNQOQLLSWGCKGRVCYTTSV 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2.6/prodata/2/iaa/5A COMB.pep.\*  
2: /cgn2.6/prodata/2/iaa/5B COMB.pep.\*  
3: /cgn2.6/prodata/2/iaa/6A COMB.pep.\*  
4: /cgn2.6/prodata/2/iaa/6B COMB.pep.\*  
5: /cgn2.6/prodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2.6/prodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	85.2	33	US-09-433-428D-6	Sequence 6, Appli
2	127	85.2	33	US-09-433-428D-25	Sequence 25, Appl
3	127	85.2	40	US-08-894-699-39	Sequence 39, Appl
4	127	85.2	40	US-08-894-699-68	Sequence 68, Appl
5	127	85.2	40	US-09-444-410-39	Sequence 39, Appl
6	127	85.2	40	US-09-444-410-68	Sequence 68, Appl
7	125	83.9	33	US-09-433-428D-30	Sequence 30, Appl
8	125	83.9	40	US-08-894-699-36	Sequence 36, Appl
9	125	83.9	40	US-09-444-410-36	Sequence 36, Appl
10	121	81.2	33	US-09-433-428D-4	Sequence 4, Appli
11	121	81.2	33	US-09-433-428D-7	Sequence 7, Appli
12	121	81.2	33	US-09-433-428D-15	Sequence 15, Appl
13	121	81.2	41	US-08-894-699-69	Sequence 69, Appl
14	121	81.2	41	US-09-444-410-69	Sequence 69, Appl
15	121	81.2	42	US-08-894-699-66	Sequence 66, Appl
16	121	81.2	42	US-09-444-410-66	Sequence 66, Appl
17	121	81.2	116	US-09-462-917A-20	Sequence 20, Appl
18	121	81.2	117	US-09-462-917A-6	Sequence 6, Appli
19	121	81.2	715	US-09-462-917A-134	Sequence 134, App
20	120	80.5	33	US-09-433-428D-5	Sequence 5, Appli
21	120	80.5	33	US-09-433-428D-8	Sequence 8, Appli
22	120	80.5	33	US-09-433-428D-16	Sequence 16, Appl
23	120	80.5	41	US-08-894-699-67	Sequence 67, Appl
24	120	80.5	41	US-09-444-410-67	Sequence 67, Appl
25	120	80.5	113	US-09-462-917A-28	Sequence 28, Appl
26	120	80.5	115	US-09-462-917A-24	Sequence 24, Appl
27	120	80.5	215	US-08-912-129A-58	Sequence 58, Appl

28	120	80.5	245	2	US-08-912-129A-48	Sequence 48, Appl
29	120	80.5	373	2	US-08-912-129A-52	Sequence 52, Appl
30	120	80.5	460	2	US-08-912-129A-60	Sequence 60, Appl
31	120	80.5	490	2	US-08-912-129A-50	Sequence 50, Appl
32	120	80.5	618	2	US-08-912-129A-54	Sequence 54, Appl
33	120	80.5	873	2	US-08-912-129A-61	Sequence 61, Appl
34	119	79.9	33	3	US-09-433-428D-10	Sequence 10, Appl
35	119	79.9	110	4	US-09-462-917A-14	Sequence 14, Appl
36	119	79.9	110	4	US-09-462-917A-16	Sequence 16, Appl
37	119	79.9	113	4	US-09-462-917A-2	Sequence 2, Appli
38	119	79.9	113	4	US-09-462-917A-10	Sequence 10, Appl
39	119	79.9	113	4	US-09-462-917A-12	Sequence 12, Appl
40	119	79.9	875	4	US-09-462-917A-18	Sequence 18, Appl
41	118.5	79.5	875	4	US-09-206-551-14	Sequence 14, Appl
42	118	79.2	33	3	US-09-433-428D-21	Sequence 21, Appl
43	118	79.2	33	3	US-09-433-428D-27	Sequence 27, Appl
44	118	79.2	40	3	US-08-894-699-37	Sequence 37, Appl
45	118	79.2	40	3	US-08-894-699-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1  
US-09-433-428D-6  
; Sequence 6, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-6

Query Match 85.2%; Score 127; DB 3; Length 33;  
Best Local Similarity 85.2%; Pred. No. 2.6e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALETLLNQOQLLSWGCKGRVCYTTSV 28  
Db 4 ALETLLNQOQLLSWGCKGRVCYTTSV 30  
|||||: ||||||||||||||||  
|||||: ||||||||||||||||

RESULT 2  
US-09-433-428D-25  
; Sequence 25, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-25

Query Match 85.2%; Score 127; DB 3; Length 33;  
Best Local Similarity 85.2%; Pred. No. 2.6e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRVCYTSV 28  
|||||:|:|||||  
Db 4 ALETLNQORLLNSWGCKGRVCYTSV 30

## RESULT 3

US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSES: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-39

Query Match 85.2%; Score 127; DB 3; Length 40;  
Best Local Similarity 85.2%; Pred. No. 3.3e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRVCYTSV 28  
|||||:|:|||||  
Db 9 ALETLNQORLLNSWGCKGRVCYTSV 35

## RESULT 4

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSES: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-68

Query Match 85.2%; Score 127; DB 3; Length 40;  
Best Local Similarity 85.2%; Pred. No. 3.3e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRVCYTSV 28  
|||||:|:|||||  
Db 9 ALETLNQORLLNSWGCKGRVCYTSV 35

## RESULT 5

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSES: P.C.

STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 85.2%; Score 127; DB 3; Length 40;  
Best Local Similarity 85.2%; Pred. No. 3.3e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQQLNSWGCKGRLVCYTSV 28  
|||||:|||||  
DB 9 ALETLNQQLNSWGCKGRLVCYTSV 35

RESULT 6  
US-09-444-410-68  
Sequence 68, Application US/09444410  
Patent No. 6270975  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 85.2%; Score 127; DB 3; Length 40;  
Best Local Similarity 85.2%; Pred. No. 3.3e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQQLNSWGCKGRLVCYTSV 28  
|||||:|||||  
DB 9 ALETLNQQLNSWGCKGRLVCYTSV 35

RESULT 7  
US-09-433-428D-30  
Sequence 30, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
APPLICANT: Zheng, Jian  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433,428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 30  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-30

Query Match 83.9%; Score 125; DB 3; Length 33;  
Best Local Similarity 81.5%; Pred. No. 5.3e-12;  
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQQLNSWGCKGRLVCYTSV 28  
|||||:|||||  
DB 4 ALETLNQQLNSWGCKGRLVCYTSV 30

RESULT 8  
US-08-894-699-36  
Sequence 36, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE

; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSES: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSES: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-699-36
;
; Query Match 83.9%; Score 125; DB 3; Length 40;
; Best Local Similarity 81.5%; Pred. No. 6.5e-12;
; Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
;
; QY 2 ALETLNQRLNSWGCKGRIVCYTSV 28
; Db 9 ALETLNQRLNSWGCKGRIVCYTSV 35
;
; RESULT 9
; US-09-444-410-36
; Sequence 36, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSES: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSES: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-444-410-36
;
; Query Match 83.9%; Score 125; DB 3; Length 40;
; Best Local Similarity 81.5%; Pred. No. 6.5e-12;
; Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
;
; QY 2 ALETLNQRLNSWGCKGRIVCYTSV 28
; Db 9 ALETLNQRLNSWGCKGRIVCYTSV 35
;
; RESULT 10
; US-09-433-428D-4
; Sequence 4, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; US-09-433-428D-4
;
; Query Match 81.2%; Score 121; DB 3; Length 33;
; Best Local Similarity 81.5%; Pred. No. 2.1e-11;
; Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
;
; QY 2 ALETLNQRLNSWGCKGRIVCYTSV 28
; Db 4 ALETLNQRLNSWGCKGRIVCYTSV 30
;
; RESULT 11
; US-09-433-428D-7
; Sequence 7, Application US/09433428D
; Patent No. 6149910



```

; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-7

Query Match      81.2%; Score 121; DB 3; Length 33;
Best Local Similarity 81.5%; Pred. No. 2.1e-11;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALETLNQRLNSWGCKGRVCYTSV 28
Db      4 ALETLNQRLNSWGCKGRVCYTSV 30

RESULT 12
US-09-433-428D-15
; Sequence 15, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-15

Query Match      81.2%; Score 121; DB 3; Length 33;
Best Local Similarity 81.5%; Pred. No. 2.1e-11;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALETLNQRLNSWGCKGRVCYTSV 28
Db      4 ALETLNQRLNSWGCKGRVCYTSV 30

RESULT 13
US-08-894-699-69
; Sequence 69, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; TITLE OF INVENTION: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

```

```

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-699-69

Query Match      81.2%; Score 121; DB 3; Length 41;
Best Local Similarity 81.5%; Pred. No. 2.7e-11;
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALETLNQRLNSWGCKGRVCYTSV 28
Db      9 ALETLNQRLNSWGCKGRVCYTSV 35

RESULT 14
US-09-444-410-69
; Sequence 69, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-69

Query Match 81.2%; Score 121; DB 3; Length 41;  
Best Local Similarity 81.5%; Pred. No. 2.7e-11;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLNQQLNSWGCKGRVYTSV 28  
|||||:|||||  
Db 9 ALETLNQQLNSWGCKGRVYTSV 35

## RESULT 15

US-08-894-699-66  
Sequence 66, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,699  
FILING DATE: 01-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR96/00294  
FILING DATE: 26-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-699-66

Query Match 81.2%; Score 121; DB 3; Length 42;  
Best Local Similarity 81.5%; Pred. No. 2.7e-11;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLNQQLNSWGCKGRVYTSV 28  
|||||:|||||  
Db 9 ALETLNQQLNSWGCKGRVYTSV 35

Search completed: May 7, 2004, 18:36:06  
Job time : 17.5522 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:34:27 ; Search time 44.7164 Seconds  
(without alignments)  
173.803 Million cell updates/sec

Title: US-09-147-362A-15  
Perfect score: 149  
Sequence: 1 RALETLLNQORLLNSWGCKRLVCYTSV 28

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121	81.2	116	14	US-10-320-786-20
2	121	81.2	117	14	US-10-320-786-6
3	121	81.2	715	14	US-10-320-786-134
4	120	80.5	113	14	US-10-320-786-28
5	120	80.5	115	14	US-10-320-786-24
6	120	80.5	215	8	US-08-911-824-58
7	120	80.5	245	8	US-08-911-824-48
8	120	80.5	281	8	US-08-911-824-120
9	120	80.5	373	8	US-08-911-824-52
10	120	80.5	460	8	US-08-911-824-60
11	120	80.5	488	8	US-08-911-824-95
12	120	80.5	490	8	US-08-911-824-50
13	120	80.5	526	8	US-08-911-824-97
14	120	80.5	618	8	US-08-911-824-54
15	120	80.5	706	8	US-08-911-824-93

Sequence 91, Appl  
Sequence 61, Appl  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 12, Appl  
Sequence 18, Appl  
Sequence 14, Appl  
Sequence 12, Appl  
Sequence 40, Appl  
Sequence 30, Appl  
Sequence 86, Appl  
Sequence 104, Appl  
Sequence 15, Appl  
Sequence 32, Appl  
Sequence 62, Appl  
Sequence 62, Appl  
Sequence 62, Appl  
Sequence 62, Appl  
Sequence 88, Appl  
Sequence 30, Appl  
Sequence 36, Appl  
Sequence 10, Appl  
Sequence 105, Appl  
Sequence 46, Appl  
Sequence 46, Appl  
Sequence 46, Appl

736 8 US-08-911-824-91  
873 8 US-08-911-824-61  
110 14 US-10-320-786-14  
110 14 US-10-320-786-16  
113 14 US-10-320-786-2  
113 14 US-10-320-786-10  
113 14 US-10-320-786-12  
113 14 US-10-320-786-18  
875 14 US-10-369-294-14  
356 14 US-10-357-400-12  
113 14 US-10-320-786-40  
32 14 US-10-364-360-30  
37 14 US-10-026-741-86  
200 9 US-09-854-816-104  
862 14 US-10-369-294-15  
32 14 US-10-364-360-32  
35 9 US-09-886-156-62  
35 9 US-09-886-150-62  
35 10 US-09-886-149-62  
35 10 US-09-886-159-62  
35 14 US-10-326-090-62  
37 14 US-10-026-741-88  
110 14 US-10-320-786-30  
110 14 US-10-320-786-36  
146 13 US-10-000-321-10  
204 9 US-09-854-816-105  
351 9 US-09-886-156-46  
351 9 US-09-886-150-46  
351 10 US-09-886-149-46  
351 10 US-09-886-159-46

## ALIGNMENTS

RESULT 1  
US-10-320-786-20  
; Sequence 20, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11382.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-20

Query Match 81.2%; Score 121; DB 14; Length 116;  
Best Local Similarity 81.5%; Pred. No. 1.7e-09;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALETLLNQORLLNSWGCKRLVCYTSV 28  
Db 29 ALETLLNQORLLNSWGCKRLVCYTSV 55

RESULT 2  
US-10-320-786-6  
; Sequence 6, Application US/10320786  
; Publication No. US20030180759A1

; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-6

Query Match 81.2%; Score 121; DB 14; Length 117;  
Best Local Similarity 81.5%; Pred. No. 1.8e-09;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ALETLNQORLNSWGCKGRLVCYTSV 28  
Db 29 ALETLNQORLNSWGCKGRLVCYTSV 55

## RESULT 3

US-10-320-786-134  
; Sequence 134, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 134  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-134

Query Match 81.2%; Score 121; DB 14; Length 715;  
Best Local Similarity 81.5%; Pred. No. 1.1e-08;  
Matches 22; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ALETLNQORLNSWGCKGRLVCYTSV 28  
Db 618 ALETLNQORLNSWGCKGRLVCYTSV 644

## RESULT 4

US-10-320-786-28  
; Sequence 28, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen

; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-28

Query Match 80.5%; Score 120; DB 14; Length 113;  
Best Local Similarity 77.8%; Pred. No. 2.4e-09;  
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ALETLNQORLNSWGCKGRLVCYTSV 28  
Db 29 ALETLNQORLNSWGCKGRLVCYTSV 55

## RESULT 5

US-10-320-786-24  
; Sequence 24, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-24

Query Match 80.5%; Score 120; DB 14; Length 115;  
Best Local Similarity 77.8%; Pred. No. 2.4e-09;  
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ALETLNQORLNSWGCKGRLVCYTSV 28  
Db 29 ALETLNQORLNSWGCKGRLVCYTSV 55

## RESULT 6

US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Roshil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

```
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL
US-08-911-824-58
```

```
Query Match      80.5%; Score 120; DB 8; Length 215;
Best Local Similarity 77.8%; Pred. No. 4.5e-09;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQQLLNLSWGCKGRLCYTSV 28
    |||||:|:|||||:|||||:|||||
DB 117 ALETLNQQLLNLSWGCKGRLCYTSV 143
```

## RESULT 7

```
US-08-911-824-48
; Sequence 48, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL
US-08-911-824-48
```

```
Query Match      80.5%; Score 120; DB 8; Length 245;
Best Local Similarity 77.8%; Pred. No. 5.1e-09;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQQLLNLSWGCKGRLCYTSV 28
    |||||:|:|||||:|||||:|||||
DB 117 ALETLNQQLLNLSWGCKGRLCYTSV 143
```

## RESULT 8

```
US-08-911-824-120
; Sequence 120, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
```

```
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 120
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL
US-08-911-824-120
```

```
Query Match      80.5%; Score 120; DB 8; Length 281;
Best Local Similarity 77.8%; Pred. No. 5.9e-09;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQQLLNLSWGCKGRLCYTSV 28
    |||||:|:|||||:|||||:|||||
DB 117 ALETLNQQLLNLSWGCKGRLCYTSV 143
```

## RESULT 9

```
US-08-911-824-52
; Sequence 52, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL
US-08-911-824-52
```

```
Query Match      80.5%; Score 120; DB 8; Length 373;
Best Local Similarity 77.8%; Pred. No. 7.8e-09;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQQLLNLSWGCKGRLCYTSV 28
    |||||:|:|||||:|||||:|||||
DB 117 ALETLNQQLLNLSWGCKGRLCYTSV 143
```

## RESULT 10

```
US-08-911-824-60
; Sequence 60, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
```

```
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pCO-8CKS
US-08-911-824-60
```

```
Query Match      80.5%; Score 120; DB 8; Length 460;
Best Local Similarity 77.8%; Pred. No. 9.6e-09;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQORLLNSWGCKGRLCYTSV 28
   |||||:|:|||||:|||||:|||||
DB 362 ALETLNQQLNLWGCKGRLCYTSV 388
```

```
RESULT 11
US-08-911-824-95
; Sequence 95, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pCO-14PL
US-08-911-824-95
```

```
Query Match      80.5%; Score 120; DB 8; Length 488;
Best Local Similarity 77.8%; Pred. No. 1e-08;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQORLLNSWGCKGRLCYTSV 28
   |||||:|:|||||:|||||:|||||
DB 117 ALETLNQQLNLWGCKGRLCYTSV 143
```

```
RESULT 12
US-08-911-824-50
; Sequence 50, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
```

```
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-50
```

```
Query Match      80.5%; Score 120; DB 8; Length 490;
Best Local Similarity 77.8%; Pred. No. 1e-08;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQORLLNSWGCKGRLCYTSV 28
   |||||:|:|||||:|||||:|||||
DB 362 ALETLNQQLNLWGCKGRLCYTSV 388
```

```
RESULT 13
US-08-911-824-97
; Sequence 97, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS
US-08-911-824-97
```

```
Query Match      80.5%; Score 120; DB 8; Length 526;
Best Local Similarity 77.8%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 ALETLNQORLLNSWGCKGRLCYTSV 28
   |||||:|:|||||:|||||:|||||
DB 362 ALETLNQQLNLWGCKGRLCYTSV 388
```

```
RESULT 14
US-08-911-824-54
; Sequence 54, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

Query Match      80.5%; Score 120; DB 8; Length 618;
Best Local Similarity 77.8%; Pred. No. 1.3e-08;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALETLNQQLRLNSWGCKGRVCYTSV 28
Db      362 ALETLNQQLRLNSWGCKGRVCYTSV 388
      |||||: |||||: |||||: |||||:
      |||||: |||||: |||||: |||||:

RESULT 15
US-08-911-824-93
; Sequence 93, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-13CKS
US-08-911-824-93

Query Match      80.5%; Score 120; DB 8; Length 706;
Best Local Similarity 77.8%; Pred. No. 1.5e-08;
Matches 21; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALETLNQQLRLNSWGCKGRVCYTSV 28
Db      608 ALETLNQQLRLNSWGCKGRVCYTSV 634
      |||||: |||||: |||||: |||||:
      |||||: |||||: |||||: |||||:

Search completed: May 7, 2004, 18:47:36
Job time : 44.7164 secs
```

THE PACE OF LIFE (1950)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:23:42 ; Search time 14.209 Seconds  
(without alignments)  
189.554 Million cell updates/sec

Title: US-09-147-362A-15  
Perfect score: 149  
Sequence: 1 RALETLNQORLLNSWGCKGRLVCYTSV 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	115	77.2	863	2 A53034	gag polyprotein -
2	110	73.8	877	2 S49197	env polyprotein p
3	109	73.2	104	2 S52930	GP41 ENV protein -
4	85	57.0	357	2 S21990	envelope protein g
5	85	57.0	358	2 S22002	envelope protein g
6	85	57.0	454	2 B41621	env polyprotein D
7	85	57.0	854	1 VCLJST	env polyprotein pr
8	85	57.0	855	2 A45713	Env transmembrane
9	84	56.4	358	2 S22000	envelope protein g
10	84	56.4	358	2 S70417	envelope protein g
11	82	55.0	357	2 S22006	envelope protein g
12	82	55.0	357	2 S21994	envelope protein g
13	82	55.0	357	2 S22004	envelope protein g
14	82	55.0	357	2 S21996	envelope protein g
15	82	55.0	357	2 S21992	envelope protein g
16	82	55.0	358	2 S21998	envelope protein g
17	82	55.0	443	2 C41621	env polyprotein P
18	82	55.0	445	2 A41621	env polyprotein M
19	82	55.0	843	1 H44001	env polyprotein pr
20	82	55.0	847	2 T09448	envelope glycoprot
21	82	55.0	847	2 S13289	env protein - huma
22	82	55.0	852	1 VCLJBR	env polyprotein -
23	82	55.0	852	2 T12016	envelope glycoprot
24	82	55.0	853	2 S54384	envelope polyprote
25	82	55.0	854	2 S13288	env protein - huma
26	82	55.0	855	1 VCLJAJ2	env polyprotein pr
27	82	55.0	855	1 VCLJZR	env polyprotein pr
28	82	55.0	856	1 VCLJH3	env polyprotein pr
29	82	55.0	856	1 VCLJVL	env polyprotein pr

30	82	55.0	856	1 VCLJ3W	env polyprotein pr
31	82	55.0	859	1 VCLJMN	env polyprotein pr
32	82	55.0	859	1 VCLJST	env polyprotein pr
33	82	55.0	861	1 VCLJLV	env polyprotein pr
34	82	55.0	861	1 VCLJSC	env polyprotein pr
35	82	55.0	868	1 VCLJH4	env polyprotein -
36	79	53.0	151	2 S30458	env protein - huma
37	79	53.0	151	2 S30459	env protein - huma
38	79	53.0	151	2 S30448	env protein - huma
39	79	53.0	151	2 S30453	env protein - huma
40	79	53.0	151	2 S30452	env protein - huma
41	79	53.0	151	2 S30450	env protein - huma
42	79	53.0	151	2 S30451	env protein - huma
43	79	53.0	151	2 S30457	env protein - huma
44	79	53.0	151	2 S30456	env protein - huma
45	79	53.0	151	2 S30455	env protein - huma

ALIGNMENTS

RESULT 1

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborght, B.; van der Groen,

J. Virol. 68, 1586-1596, 1994

A;Title: Genomic cloning and complete sequence analysis of a highly divergent African

A;Reference number: A53034; MUID:94149849; PMID:8107220

A;Accession: A53034

A;Status: preliminary

A;Molecule type: Genomic RNA

A;Residues: 1-863 <VAN>

A;Cross-references: GB:102587

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: polyprotein

Query Match

Best Local Similarity 77.2%; Score 115; DB 2; Length 863;

Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALETLNQORLLNSWGCKGRLVCYTSV 28

||||| :||| :||| :|||

Db 584 ALETLNQORLLNSWGCKGRLVCYTSV 610

RESULT 2

S49197

envelope protein precursor - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 19-Mar-1997 #sequence\_revision 24-Jul-1997 #text\_change 26-Aug-1999

C:Accession: S49197

R:Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy,

submitted to the EMBL Data Library, July 1994

A;Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: def

A;Reference number: S49197

A;Accession: S49197

A;Molecule type: DNA

A;Residues: 1-877 <CHA>

A;Cross-references: EMBL:X80020; NID:G510516; PIDN:CAA56323.1; PID:G510517

A;Experimental source: isolate VAU

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein

F;1-30/Domain: signal sequence #status predicted <SIG>

F;31-535/Product: coat protein gp120 #status predicted <CP1>

F;536-877/Product: coat protein gp41 #status predicted <CP2>

F;698-716/Domain: transmembrane #status predicted <TM>

F;59,88,139,148,159,184,189,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404

Query Match

Best Local Similarity 73.8%; Score 110; DB 2; Length 877;

Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ALETTLNQORLLNSWGCKGRVCYTSV 28  
||||| : ||||| ||||| ||||| |||||  
DB 597 ALETTIQOQLLNWGCCKNRLLCYTSV 623

## RESULT 3

S22930  
Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999  
A:Accession: S52930  
R:Cohen, J.H.M.; Guetard, D.; Philbert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de  
submitted to the EMBL Data Library, January 1995  
A:Description: A novel HIV1-O strain illustrates the diversity of the O group.  
A:Reference number: S52929  
A:Accession: S52930  
A:Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-104 <COH>  
A:Cross-references: EMBL:X84328; NID:G695526; PIDN:CAA59066.1; PID:G695527  
C:Superfamily: type E retrovirus env polyprotein

Query Match 73.2%; Score 109; DB 2; Length 104;  
Best Local Similarity 66.7%; Pred. No. 4.9e-09;  
Matches 18; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALETTLNQORLLNSWGCKGRVCYTSV 28  
||||| : ||||| ||||| ||||| |||||  
DB 35 ALETTMQOQLLNWGCCKGKCAICYTSV 61

## RESULT 4

S21990  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 20  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
A:Accession: S21990; S70423  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi  
A:Reference number: S21990  
A:Accession: S21990  
A:Molecule type: DNA  
A:Residues: 1-357 <STEL>  
A:Cross-references: EMBL:X61357; NID:G60175; PIDN:CAA43626.1; PID:G60176  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70423  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-332, 'X', 334-357 <STE2>  
A:Cross-references: EMBL:X61357; NID:G60175; PIDN:CAA43626.1; PID:G60176  
C:Superfamily: type E retrovirus env polyprotein

Query Match 57.0%; Score 85; DB 2; Length 357;  
Best Local Similarity 55.6%; Pred. No. 6.4e-05;  
Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 ALETTLNQORLLNSWGCKGRVCYTSV 28  
||||| : ||||| ||||| ||||| |||||  
DB 83 AVERYLKDOQLLNGCGSRLLCTTAV 109

## RESULT 5

S22002  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22002; S70418  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi  
A:Reference number: S21990  
A:Accession: S22002  
A:Molecule type: DNA  
A:Residues: 1-358 <STEL>  
A:Cross-references: EMBL:X61352; NID:G60186; PIDN:CAA43616.1; PID:G60187  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333, 'X', 335-358 <STE2>  
A:Cross-references: EMBL:X61352; NID:G60186  
C:Superfamily: type E retrovirus env polyprotein

Query Match 57.0%; Score 85; DB 2; Length 358;  
Best Local Similarity 55.6%; Pred. No. 6.4e-05;  
Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 ALETTLNQORLLNSWGCKGRVCYTSV 28  
||||| : ||||| ||||| ||||| |||||  
DB 84 AVERYLKDOQLLNGCGSRLLCTTAV 110

## RESULT 6

B41621  
env polyprotein D - human immunodeficiency virus type 1 (fragment)  
N:Alternate names: coat polyprotein  
N:Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 31-Jan-1997  
A:Accession: B41621  
R:Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991  
A:Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity  
A:Reference number: A41621; MUID:92107924; PMID:1763038  
A:Accession: B41621  
A:Molecule type: DNA  
A:Residues: 1-454 <BUR>  
A:Cross-references: GB:M77279  
A:Note: this virus was isolated from the daughter  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-262/Product: coat protein gp120 (fragment) #status predicted <GP1>  
F:263-454/Product: coat protein gp41 (fragment) #status predicted <GP2>  
F:435-454/Domain: transmembrane #status predicted <TMN>  
F:9,23,36,42,48,49,78,85,102,108,132,138,142,195,210,214,362,367,376/Binding site: carbo

Query Match 57.0%; Score 85; DB 2; Length 454;  
Best Local Similarity 55.6%; Pred. No. 8.1e-05;  
Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 ALETTLNQORLLNSWGCKGRVCYTSV 28  
||||| : ||||| ||||| ||||| |||||  
DB 333 ALERYLKDOQLLNGCGSRLLCTTAV 359

## RESULT 7

VCLJST  
env polyprotein precursor - simian immunodeficiency virus SIVcpz  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: simian immunodeficiency virus SIVcpz  
A:Note: host Pan troglodytes (chimpanzee)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C;Accession: S09990  
R;Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.  
Nature 345, 356-359, 1990  
A;Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.  
A;Reference number: S09983; MUID:90259077; PMID:2188136  
A;Accession: S09990  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-854 <HUE>  
A;Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36407.1; PID:g58874  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein  
F;1-30/Domain: signal sequence #status predicted <SIG>  
F;31-500/Product: coat protein gp120 #status predicted <CP1>  
F;501-854/Product: coat protein gp41 #status predicted <CP2>  
F;501-517/Domain: transmembrane #status predicted <TM1>  
F;575-693/Domain: transmembrane #status predicted <TM2>  
F;805-821/Domain: transmembrane #status predicted <TM3>  
F;134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,458  
Query Match 57.0%; Score 85; DB 1; Length 854;  
Best Local Similarity 51.3%; Pred. No. 0.00015;  
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28  
DB 572 AVERYLQDQQLGLWGCGKAVCYTTV 598  
RESULT 8  
A45713  
Env transmembrane protein gp43 - human immunodeficiency virus type 2  
C;Species: human immunodeficiency virus type 2, HIV-2  
C;Date: 04-Mar-1994 #sequence\_revision 18-Nov-1994 #text\_change 12-Apr-1995  
C;Accession: A45713  
R;Barnett, S.W.; Quiroga, M.; Werner, A.; Dina, D.; Levy, J.A.  
J. Virol. 67, 1006-1014, 1993  
A;Title: Distinguishing features of an infectious molecular clone of the highly divergent  
A;Reference number: A45713; MUID:93124535; PMID:8419635  
A;Accession: A45713  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-855 <BAR>  
A;Experimental source: UC1  
A;Note: sequence extracted from NCBI backbone (NCBI:P:122362)  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 57.0%; Score 85; DB 2; Length 855;  
Best Local Similarity 59.3%; Pred. No. 0.00015;  
Matches 16; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28  
DB 584 ATEKYLQDQALLNSWGCAFRQVCHTTV 610  
RESULT 9  
S22000  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C;Accession: S22000  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined  
A;Reference number: S21990  
A;Accession: S22000  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <STE>

A;Cross-references: EMBL:X61351  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 56.4%; Score 84; DB 2; Length 358;  
Best Local Similarity 55.6%; Pred. No. 9e-05;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28  
DB 84 AVERYLKDQQLGLWGCGRLICTTV 110  
RESULT 10  
S70417  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 3B) (fragmented)  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: patient 3B  
C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 26-Aug-1999  
C;Accession: S70417  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined  
A;Reference number: S70417; MUID:92144209; PMID:1736940  
A;Accession: S70417  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <STE>  
A;Cross-references: EMBL:X61351; NID:g60184; PIDN:CAA43614.1; PID:g60185  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 56.4%; Score 84; DB 2; Length 358;  
Best Local Similarity 55.6%; Pred. No. 9e-05;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28  
DB 84 AVERYLKDQQLGLWGCGRLICTTV 110  
RESULT 11  
S22006  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C;Accession: S70420; S22006  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined  
A;Reference number: S70417; MUID:92144209; PMID:1736940  
A;Accession: S70420  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-357 <ST2>  
A;Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191  
A;Experimental source: patient L  
A;Note: submitted to the EMBL Data Library, July 1991  
C;Superfamily: type E retrovirus env polyprotein  
Query Match 55.0%; Score 82; DB 2; Length 357;  
Best Local Similarity 51.9%; Pred. No. 0.00018;  
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28  
DB 83 AVERYLKDQQLGLWGCGRLICTTAV 109  
RESULT 12  
S21994  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: isolate 27B  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:53:36 ; Search time 10.0299 Seconds  
(without alignments)  
145.363 Million cell updates/sec

Title: US-09-147-362A-15  
Perfect score: 149  
Sequence: 1 RAETLLNQRLNSWGCKGRVLCYTSTV 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	57.0	854	1 ENV_SIVCZ	P17281 chimpanzee
2	84	56.4	857	1 ENV_HV2KR	Q74126 human immun
3	82	55.0	843	1 ENV_HV1Y2	P35961 human immun
4	82	55.0	847	1 ENV_HV1S1	P19550 human immun
5	82	55.0	847	1 ENV_HV1W2	P05880 human immun
6	82	55.0	851	1 ENV_HV1B8	P04582 human immun
7	82	55.0	852	1 ENV_HV1BN	P12488 human immun
8	82	55.0	852	1 ENV_HV1S3	P19549 human immun
9	82	55.0	853	1 ENV_HV1MF	P19551 human immun
10	82	55.0	853	1 ENV_HV1Z2	P12487 human immun
11	82	55.0	855	1 ENV_HV1A2	P03378 human immun
12	82	55.0	855	1 ENV_HV1OY	P20888 human immun
13	82	55.0	855	1 ENV_HV1Z6	P04580 human immun
14	82	55.0	856	1 ENV_HV1B1	P03375 human immun
15	82	55.0	856	1 ENV_HV1H2	P04578 human immun
16	82	55.0	856	1 ENV_HV1LW	Q70626 human immun
17	82	55.0	856	1 ENV_HV1MN	P05877 human immun
18	82	55.0	856	1 ENV_HV1PV	P03376 human immun
19	82	55.0	856	1 ENV_HV1SC	P05878 human immun
20	82	55.0	856	1 ENV_HV1W1	P31872 human immun
21	82	55.0	856	1 ENV_HV2NZ	P05883 human immun
22	82	55.0	861	1 ENV_HV1BR	P03377 human immun
23	82	55.0	865	1 ENV_HV1RH	P04579 human immun
24	82	55.0	867	1 ENV_HV1J3	P12489 human immun
25	82	55.0	868	1 ENV_HV1C4	P05879 human immun
26	80	53.7	848	1 ENV_HV1JR	P20871 human immun
27	80	53.7	856	1 ENV_HV1H3	P04624 human immun
28	79	53.0	712	1 ENV_HV2S2	P32536 human immun
29	79	53.0	851	1 ENV_HV2D1	P17755 human immun
30	79	53.0	851	1 ENV_HV2G1	P18040 human immun
31	79	53.0	854	1 ENV_SIVAI	Q02837 simian immu
32	79	53.0	856	1 ENV_HV1ZB	P05881 human immun
33	79	53.0	858	1 ENV_HV2ZB	P04577 human immun

#### ALIGNMENTS

##### RESULT 1

ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
DE	ENV.			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1.";			
RL	Nature 345:356-359(1990).			
CC	-1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X52154; CAA36407.1; -			
DR	PIR; S09990; VCLJST.			
DR	HIV; X52154; ENV\$CPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Glycoprotein; Transmembrane;			
FT	Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .) (POTENTIAL).

P24105 human immun  
P15831 human immun  
P20872 human immun  
P18094 human immun  
P05886 simian immu  
P12492 simian immu  
P19503 simian immu  
P18799 human immun  
P12449 human immun  
P04583 human immun  
P05882 human immun  
P08810 simian immu

FT	CARBOHYD	233	233	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	264	270	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	281	281	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	391	391	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	440	440	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	455	455	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	628	628	N-LINKED (GLCNAC. .)	(POTENTIAL).
SQ	SEQUENCE	857 AA;	98689 MW; F881C6755B5746DF CRC64;		

Query Match 56.4%; Score 84; DB 1; Length 857;  
Best Local Similarity 59.3%; Pred.No. 5.1e-05;  
Matches 16; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```

QY      2 ALETLNQORLLNSWGCKRLVCVTSV 28
       |:|:|||||||
DB     573 AIEKYLRDAQRLNSWGAFRQCIVTV 599

```

RESULT 3  
ENV\_HVIY2 STANDARD; PRT; 843 AA.

ID	ENV_HVIY2	STANDARD;	PRT;	843 AA.
AC	F35961;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE	GN			
OS	Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=36377;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93021387; PubMed=1404605;			
RA	Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H., Shaw G.M.;			
RA	"Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation."			
RL	J. Virol. 66:6587-6600(1992).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenses@isb-sib.ch).			
DR	EMBL; M33258; -; NOT_ANNOTATED_CDS.			
DR	PIR; H44001; H44001.			
DR	PDB; 1G9N; 27-DEC-00.			
DR	InterPro; IPR000328; Env_GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.			
KW	SIGNAL	1	29	
FT	CHAIN	30	489	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	490	843	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	738	755	POTENTIAL.
FT	DISULFID	53	73	BY SIMILARITY.
FT	DISULFID	118	201	BY SIMILARITY.
FT	DISULFID	125	192	BY SIMILARITY.
FT	DISULFID	130	155	BY SIMILARITY.







```
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 55.0%; Score 82; DB 1; Length 851;
Best Local Similarity 51.9%; Pred. No. 0.0001;
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ALETLNQQLLNQGLNMGCKRLVCYTSV 28
Db 577 AVERYLKDQQLGNGSGKLICTTAV 603

RESULT 7
ENV_HV1BN
ID ENV_HV1BN STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Danekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21098; AAA44221.1; -.
CC PIR; A31667; VCLJBR.
CC PDB; 1IM7; 23-OCT-02.
CC HIV; M21098; ENV$BRVA.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 30
CC CHAIN 31 507
CC CHAIN 508 852
CC EXTERIOR MEMBRANE GLYCOPROTEIN.
CC TRANSMEMBRANE GLYCOPROTEIN.
CC FT
```

```
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 155 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 439 BY SIMILARITY.
FT DISULFID 383 412 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 55.0%; Score 82; DB 1; Length 852;
Best Local Similarity 51.9%; Pred. No. 0.0001;
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ALETLNQQLLNQGLNMGCKRLVCYTSV 28
Db 578 AVERYLKDQQLGNGSGKLICTTAV 604

RESULT 8
ENV_HV1S3
ID ENV_HV1S3 STANDARD; PRT; 852 AA.
AC P19549;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OX Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```





```

FT DISULFID 125 199 BY SIMILARITY. (POTENTIAL).
FT DISULFID 130 155 BY SIMILARITY. (POTENTIAL).
FT DISULFID 221 250 BY SIMILARITY. (POTENTIAL).
FT DISULFID 231 242 BY SIMILARITY. (POTENTIAL).
FT DISULFID 299 333 BY SIMILARITY. (POTENTIAL).
FT DISULFID 380 442 BY SIMILARITY. (POTENTIAL).
FT DISULFID 387 415 BY SIMILARITY. (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 55.0%; Score 82; DB 1; Length 855;
Best Local Similarity 51.9%; Pred. No. 0.0001;
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ALETLLNQRLNSWGCKGRLYCYTSV 28
Db 581 AVERYLRDQQLLGWCGSKLICITTV 607

RESULT 12
ENV_HV10Y STANDARD; PRT; 855 AA.
AC P20888;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR HIV; M26727; AA83397.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 30 509 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 510 855 BY SIMILARITY.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 210 BY SIMILARITY.
FT DISULFID 125 201 BY SIMILARITY.
FT DISULFID 130 162 BY SIMILARITY.
FT DISULFID 223 252 BY SIMILARITY.
FT DISULFID 233 244 BY SIMILARITY.
FT DISULFID 301 335 BY SIMILARITY.
FT DISULFID 381 442 BY SIMILARITY.
FT DISULFID 388 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 55.0%; Score 82; DB 1; Length 855;
Best Local Similarity 51.9%; Pred. No. 0.0001;
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ALETLLNQRLNSWGCKGRLYCYTSV 28
Db 581 AVERYLRDQQLLGWCGSKLICITTV 607

RESULT 13
ENV_HV126 STANDARD; PRT; 855 AA.
ID ENV_HV126
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
```

```

FT CARBOHYD      615      615      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     624      624      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     636      636      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     673      673      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE      855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;

Query Match          55.0%; Score 82; DB 1; Length 855;
Best Local Similarity 51.9%; Pred. No. 0.0001;
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      2 ALETLINQORLLNSWGCKRLVCYTSS 28
       | | | | | | | | | | | | | | |
Db      581 AVERYLKDOQLLGICGSGKLICTTV 607

RESULT 14
ENV_HV1B1
ID ENV HV1B1 STANDARD; PRT; 856 AA.
AC F03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane
DE glycoprotein (gp120); Transmembrane glycoprotein (gp41)].
EN
GN Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11678;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L.J., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Skaal F.;
RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RT Nature 313:277-284 (1985).
RL [2]
RN DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=90285159; PubMed=2355006;
RX Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells.";
RT J. Biol. Chem. 265:10373-10382 (1990).
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15654; AAA44205.1; -.
DR FIR; A03973; VCLJH3.
DR HIV; M15654; ENVSBH102.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
KW SIGNAL.
FT CHAIN 1 30
FT CHAIN 512 856 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 54 74 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
```





THIS PAGE BLANK (00010)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:22:06 ; Search time 41.791 Seconds  
(without alignments)  
211.397 Million cell updates/sec

Title: US-09-147-362A-15

Perfect score: 149

Sequence: 1 RALETLLNQORLLNSWGCKGRVLCYTSTV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	85.2	116	15 Q7ZJN9	Q7zjn9 human immun
2	127	85.2	216	15 Q9IEC5	Q9iec5 human immun
3	127	85.2	219	15 Q9IEB6	Q9ieb6 human immun
4	127	85.2	890	15 Q8Q7G2	Q8q7g2 human immun
5	125	83.9	219	15 Q9IEC8	Q9iec8 human immun
6	124	83.2	130	15 Q9IHU9	Q9ihu9 human immun
7	124	83.2	872	15 Q8Q7H0	Q8q7h0 human immun
8	124	83.2	882	15 Q8Q7F9	Q8q7f9 human immun
9	124	83.2	887	15 Q8Q7G9	Q8q7g9 human immun
10	123	82.6	242	15 Q9IE31	Q9ie31 human immun
11	122	81.9	216	15 Q9IEA5	Q9iea5 human immun
12	122	81.9	544	15 Q9IED9	Q9ied9 human immun
13	121	81.2	115	15 Q7ZJN8	Q7zjn8 human immun
14	121	81.2	116	15 Q40459	Q40459 human immun
15	121	81.2	118	15 Q40451	Q40451 human immun
16	121	81.2	131	15 Q9WR05	Q9wr05 human immun

17	121	81.2	131	15 Q9IHU0	Q9ihu0 human immun
18	121	81.2	137	15 Q9IHV5	Q9ihv5 human immun
19	121	81.2	213	15 Q9IEC3	Q9iec3 human immun
20	121	81.2	214	15 Q9DIK3	Q9dik3 human immun
21	121	81.2	342	15 Q11942	Q11942 human immun
22	121	81.2	532	15 Q9IEE9	Q9iee9 human immun
23	121	81.2	871	15 Q57073	Q57073 human immun
24	121	81.2	871	15 Q57074	Q57074 human immun
25	121	81.2	871	15 Q8Q7I2	Q8q7i2 human immun
26	121	81.2	872	15 Q900Y5	Q900y5 human immun
27	121	81.2	872	15 Q57072	Q57072 human immun
28	121	81.2	875	15 Q8Q7H4	Q8q7h4 human immun
29	121	81.2	900	15 Q9QN28	Q9qzn8 human immun
30	120	80.5	114	15 Q40448	Q40448 human immun
31	120	80.5	116	15 Q40458	Q40458 human immun
32	120	80.5	116	15 Q40449	Q40449 human immun
33	120	80.5	117	15 Q9Q6F2	Q9q6f2 human immun
34	120	80.5	124	15 Q9IHU7	Q9ihu7 human immun
35	120	80.5	132	15 Q9IHV0	Q9ihv0 human immun
36	120	80.5	134	15 Q9IHV4	Q9ihv4 human immun
37	120	80.5	155	15 Q8J3Q4	Q8j3q4 human immun
38	120	80.5	158	15 Q8J3N6	Q8j3n6 human immun
39	120	80.5	172	15 Q9IEB3	Q9ieb3 human immun
40	120	80.5	183	15 Q9IEC1	Q9iec1 human immun
41	120	80.5	200	15 Q9IEB8	Q9ieb8 human immun
42	120	80.5	210	15 Q9IEB7	Q9ieb7 human immun
43	120	80.5	213	15 Q9IEC4	Q9iec4 human immun
44	120	80.5	220	15 Q9IEC9	Q9iec9 human immun
45	120	80.5	234	15 Q9IEC2	Q9iec2 human immun

#### ALIGNMENTS

#### RESULT 1

Q7ZJN9 PRELIMINARY; PRT; 116 AA.  
ID Q7ZJN9  
AC Q7ZJN9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=08692A;  
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
(Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
integrase, and env gp41.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214120; AAO61840.1; -;  
DR GO; GO:0019031; C:viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 13865 MW; AFIDFPC57E059061 CRC64;

Query Match 85.2%; Score 127; DB 15; Length 116;  
Best Local Similarity 85.2%; Pred. No. 1.3e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLLNQORLLNSWGCKGRVLCYTSTV 28

Db 22 ALETLLNQORLLNSWGCKGRVLCYTSTV 48

#### RESULT 2

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5

AC Q9IEC5;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF07;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236394; CAB96243.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 216 216  
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 85.2%; Score 127; DB 15; Length 216;  
Best Local Similarity 85.2%; Pred. No. 2.5e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLLNQRLNSWGCKGRVLCYTSV 28  
|||||:|:|||||  
Db 37 ALETLLNQRLNSWGCKGRVLCYTSV 63

RESULT 3  
Q9IEB6 PRELIMINARY; PRT; 219 AA.  
AC Q9IEB6;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF14;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236404; CAB96252.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 85.2%; Score 127; DB 15; Length 219;  
Best Local Similarity 85.2%; Pred. No. 2.6e-12;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLLNQRLNSWGCKGRVLCYTSV 28  
|||||:|:|||||  
Db 47 ALETLLNQRLNSWGCKGRVLCYTSV 73

RESULT 4  
Q8Q7G2 PRELIMINARY; PRT; 890 AA.  
AC Q8Q7G2;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97US08692A;  
RX MEDLINE=21849375; PubMed=11860674;  
RA Yanaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
RT Phylogenetic Clusters."  
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
DR EMBL; AF383259; AAL98881.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 890 AA; 100610 MW; BBE816BC5C45EE23 CRC64;

Query Match 85.2%; Score 127; DB 15; Length 890;  
Best Local Similarity 85.2%; Pred. No. 1.1e-11;  
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLLNQRLNSWGCKGRVLCYTSV 28  
|||||:|:|||||  
Db 610 ALETLLNQRLNSWGCKGRVLCYTSV 636

RESULT 5  
Q9IEC8 PRELIMINARY; PRT; 219 AA.  
AC Q9IEC8;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF02;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236391; CAB96240.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;

```
Query Match      83.9%; Score 125; DB 15; Length 219;
Best Local Similarity 81.5%; Pred. No. 5.4e-12;
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28
DB 46 ALETLNQORLLNSWGCKGRLVCYTSV 72

RESULT 6
Q91HU9
ID Q91HU9 PRELIMINARY; PRT; 130 AA.
AC Q91HU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CH798;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjuno P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229235; AAF71912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 130
FT NON_TER 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match      83.2%; Score 124; DB 15; Length 130;
Best Local Similarity 81.5%; Pred. No. 4.5e-12;
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28
DB 30 ALETLNQORLLNSWGCKGRLVCYTSV 56

RESULT 7
Q8Q7H0
ID Q8Q7H0 PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CWA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383251; AAL98873.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 627BCA5D0C09C6D4 CRC64;

Query Match      83.2%; Score 124; DB 15; Length 882;
Best Local Similarity 81.5%; Pred. No. 3.4e-11;
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQORLLNSWGCKGRLVCYTSV 28
DB 603 ALETLNQORLLNSWGCKGRLVCYTAV 629

RESULT 9
Q8Q7G9
ID Q8Q7G9 PRELIMINARY; PRT; 887 AA.
AC Q8Q7G9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
```

```

OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=58CMABBI41;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383252; AAL98874.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 887 AA; 99559 MW; 88954F4ED76A314A CRC64;

Query Match 83.2%; Score 124; DB 15; Length 887;
Best Local Similarity 81.5%; Pred. No. 3.5e-11;
Matches 22; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQRLNSGCKGRLVCYTSV 28
DB 608 ALETLIQQLNSGCKGKLCYTSV 634

RESULT 10
Q9IE31
ID Q9IE31 PRELIMINARY; PRT; 242 AA.
AC Q9IE31;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF112;
RA Roques P., Robertson D., Sousquiere S., Damond F., Mauclore P.,
RA Deplenne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ243366; CAB96336.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 242
SQ SEQUENCE 242 AA; 27539 MW; DE8A73DF0B8A6FD7 CRC64;

Query Match 82.6%; Score 123; DB 15; Length 242;
Best Local Similarity 77.8%; Pred. No. 1.3e-11;
Matches 21; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALETLNQRLNSGCKGRLVCYTSV 28
DB 54 ALETLIQQLNSGCKGKLCYTSV 80

RESULT 11
Q9IEA5
ID Q9IEA5 PRELIMINARY; PRT; 216 AA.

```

```

AC Q9IEA5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF26;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ236415; CAB96263.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25003 MW; 0A5AC218BFA8932 CRC64;

Query Match 81.9%; Score 122; DB 15; Length 216;
Best Local Similarity 78.6%; Pred. No. 1.6e-11;
Matches 22; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RALETLNQRLNSGCKGRLVCYTSV 28
DB 31 QALETIQQLNSGCKGRLVCYTSV 58

RESULT 12
Q9IED9
ID Q9IED9 PRELIMINARY; PRT; 544 AA.
AC Q9IED9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF26;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauclore P.,
RA Deplenne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ133068; CAB96229.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 544
SQ SEQUENCE 544 AA; 61398 MW; FC3CPA4E31DB6D50 CRC64;

Query Match 81.9%; Score 122; DB 15; Length 544;
Best Local Similarity 78.6%; Pred. No. 4.4e-11;
Matches 22; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```



THE NEW YORK PUBLIC LIBRARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:51:11 ; Search time 36.2836 Seconds  
(without alignments)  
132.382 Million cell updates/sec

Title: US-09-147-362A-16

Perfect score: 98  
Sequence: 1 RLNSWCKGRLVCYTSV 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	17	2 AAW80474	Aaw80474 Peptide d
2	93	94.9	22	2 AAW80472	Aaw80472 Peptide d
3	93	94.9	28	2 AAW80473	Aaw80473 Peptide d
4	93	94.9	33	3 AAB12231	Aab12231 Partial s
5	93	94.9	33	3 AAB12212	Aab12212 Partial s
6	93	94.9	40	2 AAW07346	Aaw07346 Partial s
7	93	94.9	40	2 AAW07352	Aaw07352 Partial s
8	91	92.9	23	3 AAB12255	Aab12255 HIV-1 gp4
9	91	92.9	23	3 AAB12254	Aab12254 HIV-1 gp4
10	91	92.9	23	3 AAB12256	Aab12256 HIV-1 gp4
11	91	92.9	28	3 AAB12257	Aab12257 HIV-1 gp4
12	91	92.9	30	3 AAB12264	Aab12264 HIV-1 gp4
13	91	92.9	33	3 AAB12236	Aab12236 Partial s
14	91	92.9	35	3 AAB12259	Aab12259 Group O H
15	91	92.9	40	2 AAW07343	Aaw07343 Partial s
16	91	92.9	149	3 AAB12262	Aab12262 HIV group
17	91	92.9	220	3 AAB12261	Aab12261 HIV group
18	91	92.9	368	3 AAB12253	Aab12253 DHFR-HENV
19	91	92.9	439	3 AAB12252	Aab12252 DHFR-HENV
20	90	91.8	32	2 AAW80469	Aaw80469 Peptide d
21	89	90.8	16	2 AAW80468	Aaw80468 Peptide d
22	89	90.8	36	2 AAY30515	Aay30515 Allelic p
23	89	90.8	36	3 AAY55786	Aay55786 HIV (Grou
24	89	90.8	36	3 AAY79777	Aay79777 HIV detec
25	89	90.8	36	3 AAY67615	Aay67615 Peptide #

26	89	90.8	36	3 AAY67613	Aay67613 Peptide #
27	89	90.8	36	4 AAB62597	Aab62597 Peptide a
28	88	89.8	22	2 AAW80461	Aaw80461 Peptide d
29	88	89.8	22	2 AAW80462	Aaw80462 Peptide d
30	88	89.8	23	2 AAR86282	Aar86282 HIV isola
31	88	89.8	23	2 AAW67027	Aaw67027 Peptide w
32	88	89.8	32	4 AAG65201	Aag65201 Human imm
33	88	89.8	33	3 AAB12208	Aab12208 Partial s
34	88	89.8	35	2 AAR51690	Aar51690 HIV-type
35	88	89.8	35	2 AAW93073	Aaw93073 HIV isola
36	88	89.8	36	2 AAY31617	Aay31617 HIV mutan
37	88	89.8	36	2 AAY55784	Aay55784 HIV gp41
38	88	89.8	36	3 AAY67607	Aay67607 Peptide #
39	88	89.8	36	3 AAY67606	Aay67606 Peptide #
40	88	89.8	36	4 AAB62574	Aab62574 HIV-1 gp4
41	88	89.8	167	2 AAW69320	Aaw69320 Anti-HIV-
42	88	89.8	173	3 AAU77260	Aau77260 Protein M
43	88	89.8	204	2 AAY22909	Aay22909 SEQ ID NO
44	88	89.8	204	5 ABG68380	Abg68380 Envelope
45	88	89.8	204	6 ABUS7787	Abu57787 Human imm

ALIGNMENTS

RESULT 1  
AAW80474  
ID AAW80474 standard; peptide; 17 AA.  
AC AAW80474;  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
PD 15-OCT-1998.  
XX  
PF 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
XX  
PR 24-FEB-1998; 98FR-00002212.  
XX  
(SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX  
PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX  
WPI; 1998-583190/49.  
XX  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
XX  
PT immune deficiency virus of group O.  
XX  
PS Claim 6; Page 45; 55pp; French.  
XX  
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX  
CC connected around short highly conserved sequences present in isolates of  
XX  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
XX  
CC immunological reagents for detecting infection by group O human immune  
XX  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX  
CC (Updated on 27-AUG-2003 to correct OS field.)  
SQ Sequence 17 AA;

Query Match 100.0%; Score 98; DB 2; Length 17;





PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX  
 PI De Leys R, Zheng J;  
 XX  
 DR WPI; 2000-402205/35.  
 XX  
 PT New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.  
 XX  
 XX Example 1; Fig 1; 52pp; English.  
 PS  
 CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain MAN. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. MAN is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 33 AA;

Query Match 94.9%; Score 93; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRVCYTSV 17  
 |||||  
 DB 15 LNSWGCKGRVCYTSV 30

RESULT 5  
 AAB12212  
 ID AAB12212 standard; peptide; 33 AA.  
 XX  
 AC AAB12212;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 10-NOV-2000 (first entry)  
 XX  
 DE Partial sequence of HIV-1 strain BCF13 gp41 immunodominant region.  
 XX  
 KW HIV-1; AIDS: human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF13.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN EP1013766-A2.  
 XX  
 PD 28-JUN-2000.  
 XX  
 PF 29-NOV-1999; 99EP-00309491.  
 XX  
 PR 30-NOV-1998; 98US-0110292P.  
 PR 08-FEB-1999; 99US-0119138P.  
 PR 04-NOV-1999; 99US-00433428.  
 XX  
 (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 PA  
 XX  
 PI De Leys R, Zheng J;  
 XX  
 DR WPI; 2000-402205/35.  
 XX  
 PT New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.  
 XX  
 XX Example 1; Fig 1; 52pp; English.  
 PS  
 CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain BCF13. HIV is the principle aetiological

CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
 CC envelope protein, and may be used as an antigen for the detection of  
 CC antibodies produced in response to HIV infection. BCF13 is a member of  
 CC HIV group O (outlier). The present sequence is the immunodominant region  
 CC of gp41. This sequence was used in a sequence homology alignment, which  
 CC in turn was used to derive a consensus sequence peptide: peptide 147  
 CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 33 AA;

Query Match 94.9%; Score 93; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRVCYTSV 17  
 |||||  
 DB 15 LNSWGCKGRVCYTSV 30

RESULT 6  
 AAW07346  
 ID AAW07346 standard; peptide; 40 AA.  
 XX  
 AC AAW07346;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95FR-00002236.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;  
 XX  
 DR WPI; 1996-412779/41.  
 DR N-PSDB; AAT44922.  
 XX  
 PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 34; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MVP5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC C2V3-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC C2V3-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the

CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
CC 2003 to standardise OS field)

XX Sequence 40 AA;

Query Match 94.9%; Score 93; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
| | | | | | | | | | | | | | | | | |  
Db 20 LNSWGCKGRLVCYTSV 35

## RESULT 7

AAW07352  
ID AAW07352 standard; peptide; 40 AA.

XX AAW07352;

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF13.

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
KW immunogen; antibody.

XX Human immunodeficiency virus 1.

OS WO9627013-A1.

XX 06-SEP-1996.

XX 26-FEB-1996; 96WO-FR000294.

XX 27-FEB-1995; 95FR-00002236.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
PT antibodies - useful for diagnosis, screening and typing, or as  
PT immunogens.

XX Claim 12; Page 46; 71pp; French.

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
CC into 2 major groups based on the nucleotide sequences of the envelop gene  
CC (env): group M containing sub-groups A-G, and group O containing the  
CC strains ANT70 and MVP5180. The invention relates to the discovery of  
CC several new strains of HIV-1 which can be placed in group O, based on the  
CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
CC and AAW07329-64). The novel strains have been deposited as retroviruses  
CC CCNCM I-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
CC the strain BCF13 and corresponds to a fragment of the gp41 protein  
CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
CC also for screening and typing of such strains. Peptides encoded by the  
CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
CC 2003 to standardise OS field)

XX Sequence 40 AA;

Query Match 94.9%; Score 93; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
| | | | | | | | | | | | | | | | | |  
Db 20 LNSWGCKGRLVCYTSV 35

## RESULT 8

AA112255  
ID AAB12255 standard; peptide; 23 AA.

XX AAB12255;

DT 12-SEP-2003 (revised)

DT 10-NOV-2000 (first entry)

XX HIV-1 gp41 immunodominant region consensus sequence peptide 147 # 2.

XX HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;  
KW acquired immunodeficiency syndrome; group O HIV; gp41.

XX Human immunodeficiency virus 1.

XX Key Location/Qualifiers

FT Misc-difference 1

FT /note= "Any natural amino acid apart from L-asparagine"

XX EP1013766-A2.

XX 28-JUN-2000.

XX 29-NOV-1999; 99EP-00309491.

XX 30-NOV-1998; 98US-0110292P.

PR 08-FEB-1999; 99US-0119138P.

PR 04-NOV-1999; 99US-00433428.

XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.

XX De Leys R, Zheng J;

XX WPI; 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for  
PT detection of antibodies produced in response to human immunodeficiency  
PT virus group O antibodies.

XX Claim 1; Page 35; 52pp; English.

XX The present sequence is a peptide 147 related peptide from Human  
CC Immunodeficiency Virus Type 1 (HIV-1). This sequence is a partial  
CC consensus sequence of the immunodominant region of gp41 protein derived  
CC from a variety of HIV-1 group O (outlier) strains: ANT70, MVP5180, VAV,  
CC DUR, POC, PAN, LOB, MAN, NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686,  
CC ABT063, ABT124, ABT123, ABT156, 193Ha, CDC7755, CDC1897, HLD28, 1515,  
CC 1516, D47-2d, HCYT2c, Nr42 and PE41 (see AAB12207 to AAB12236). HIV is  
CC the principle aetiological agent for acquired immunodeficiency syndrome  
CC (AIDS). gp41 is a HIV envelope protein, and so the present sequence may  
CC be used as an antigen for the detection of antibodies produced in  
CC response to HIV infection. (Updated on 12-SEP-2003 to standardise OS  
CC field)

XX Sequence 23 AA;

Query Match 92.9%; Score 91; DB 3; Length 23;

Best Local Similarity 87.5%; Pred. No. 2e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTS 16  
| | | | | | | | | | | | | | | | | |

DE	HIV-1 gp41 immunodominant region consensus sequence peptide 147-4.
XX	
KW	HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;
KW	acquired immunodeficiency syndrome; group O HIV; gp41.
XX	
OS	Human immunodeficiency virus 1.
XX	
FN	EP1013766-A2.
XX	
PD	28-JUN-2000.
XX	
PF	29-NOV-1999; 99EP-00309491.
XX	
PR	30-NOV-1998; 98US-0110292P.
PR	08-FEB-1999; 98US-0119138P.
PR	04-NOV-1999; 99US-00433428.
XX	
PA	(ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX	
PI	De Leys R, Zheng J;
XX	
DR	WPI; 2000-402205/35.
XX	
PT	New antigenic peptides and peptide functional derivatives, useful for
PT	detection of antibodies produced in response to human immunodeficiency
PT	virus group O antibodies.
XX	
PS	Claim 1; Page 36; 52pp; English.
XX	
CC	The present sequence is peptide 147-4 from Human Immunodeficiency Virus
CC	Type 1 (HIV-1). This sequence is a partial consensus sequence of the
CC	immunodominant region of gp41 protein derived from a variety of HIV-1
CC	group O (ourlier) strains: ANT70, MVP5180, VAU, DUR, POC, FAN LOB, MAN,
CC	NAN, ESS, NKO, BCR09, BCF13, BCF14, 686, ABT063, ABT124, ABT123,
CC	ABT2156, 193HA, CDC755, CDC1897, HLD28, 1515, 1516, D47-2d, HCYT2c, Nr42
CC	and PE41 (see AABI2207 to AABI2236). HIV is the principle aetiological
CC	agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC	envelope protein, and so the present sequence may be used as an antigen
CC	for the detection of antibodies produced in response to HIV infection.
CC	(Updated on 12-SEP-2003 to standardise OS field)
XX	
SQ	Sequence 23 AA;
	Query Match 92.9%; Score 91; DB 3; Length 23;
	Best Local Similarity 87.5%; Pred.No. 2e-06; Indels 0; Gaps 0;
	Matches 14; Conservative 2; Mismatches 0;
Qy	1 RLNSWGCKGRVCYTS 16      ::
Db	4 RLNSWGCKGRIICYTS 19      ::
RESULT 11	
AABI2257	ID ID AABI2257 standard; peptide; 28 AA.
XX	AAI12257;
AC	
XX	
DT	12-SEP-2003 (revised)
DT	10-NOV-2000 (first entry)
XX	
DE	HIV-1 gp41 immunodominant region consensus sequence peptide 147 # 3.
XX	
KW	HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;
KW	acquired immunodeficiency syndrome; group O HIV; gp41.
XX	
OS	Human immunodeficiency virus 1.
XX	
Key	Location/Qualifiers
FH	Misc-difference 6
FT	/note= "Any natural amino acid apart from L-asparagine"
FT	
DN	EP1013766-A2

```

XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX CC The present sequence is a peptide 147 related peptide from Human
CC Immunodeficiency Virus Type 1 (HIV-1). This sequence is a partial
CC consensus sequence of the immunodominant region of gp41 protein derived
CC from a variety of HIV-1 group O (outlier) strains: ANT70, MVP5180, VAU,
CC DUR, POC, FAN, LOB, MAN, NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686,
CC ABT063, ABT124, ABT1123, ABT2156, 193Ha, CDC7755, CDC1897, HLD28, 1515,
CC 1516, D47-2d, HCVT2C, N42 and PE41 (see AAB12207 to AAB12236). HIV is
CC the principle aetiological agent for acquired immunodeficiency syndrome
CC (AIDS). gp41 is a HIV envelope protein, and so the present sequence may
CC be used as an antigen for the detection of antibodies produced in
CC response to HIV infection. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX SQ Sequence 28 AA;
    Query Match 92.9%; Score 91; DB 3; Length 28;
    Best Local Similarity 87.5%; Pred. No. 2.4e-06;
    Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLNSWGCKGRUVCYTS 16
    |||||:::||||
Db 9 RLNSWGCKGRUVCYTS 24
    |||||:::||||

RESULT 12
AAB12264
ID AAB12264 standard; peptide; 30 AA.
XX AC AAB12264;
XX DT 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX DE HIV-1 gp41 immunodominant region consensus sequence peptide 147-5.
XX KW HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;
KW acquired immunodeficiency syndrome; group O HIV; gp41.
XX OS Human immunodeficiency virus 1.
XX PN EP1013766-A2.
XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX CC The present sequence is a peptide 147 related peptide from Human
CC Immunodeficiency Virus Type 1 (HIV-1). This sequence is a partial
CC consensus sequence of the immunodominant region of gp41 protein derived
CC from a variety of HIV-1 group O (outlier) strains: ANT70, MVP5180, VAU,
CC DUR, POC, FAN, LOB, MAN, NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686,
CC ABT063, ABT124, ABT1123, ABT2156, 193Ha, CDC7755, CDC1897, HLD28, 1515,
CC 1516, D47-2d, HCVT2C, N42 and PE41 (see AAB12207 to AAB12236). HIV is
CC the principle aetiological agent for acquired immunodeficiency syndrome
CC (AIDS). gp41 is a HIV envelope protein, and so the present sequence may
CC be used as an antigen for the detection of antibodies produced in
CC response to HIV infection. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX SQ Sequence 28 AA;
    Query Match 92.9%; Score 91; DB 3; Length 28;
    Best Local Similarity 87.5%; Pred. No. 2.4e-06;
    Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLNSWGCKGRUVCYTS 16
    |||||:::||||
Db 9 RLNSWGCKGRUVCYTS 24
    |||||:::||||

RESULT 13
AAB12236
ID AAB12236 standard; peptide; 33 AA.
XX AC AAB12236;
XX DT 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX DE Partial sequence of HIV-1 strain ESS gp41 immunodominant region.
XX KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; ESS.
XX OS Human immunodeficiency virus 1.
XX PN EP1013766-A2.
XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX CC The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain ESS. HIV is the principle aetiological agent

```

```

PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX CC The present sequence is peptide 147-5 from Human Immunodeficiency Virus
CC Type 1 (HIV-1). This sequence is a partial consensus sequence of the
CC immunodominant region of gp41 protein derived from a variety of HIV-1
CC group O (outlier) strains: ANT70, MVP5180, VAU, DUR, POC, FAN, LOB, MAN,
CC NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686, ABT063, ABT124, ABT1123,
CC ABT2156, 193Ha, CDC7755, CDC1897, HLD28, 1515, 1516, D47-2d, HCVT2C, N42
CC and PE41 (see AAB12207 to AAB12236). HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and so the present sequence may be used as an antigen
CC for the detection of antibodies produced in response to HIV infection.
CC (Updated on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 30 AA;
    Query Match 92.9%; Score 91; DB 3; Length 30;
    Best Local Similarity 87.5%; Pred. No. 2.6e-06;
    Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLNSWGCKGRUVCYTS 16
    |||||:::||||
Db 11 RLNSWGCKGRUVCYTS 26
    |||||:::||||

RESULT 13
AAB12236
ID AAB12236 standard; peptide; 33 AA.
XX AC AAB12236;
XX DT 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX DE Partial sequence of HIV-1 strain ESS gp41 immunodominant region.
XX KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; ESS.
XX OS Human immunodeficiency virus 1.
XX PN EP1013766-A2.
XX PD 28-JUN-2000.
XX PF 29-NOV-1999; 99EP-00309491.
XX PR 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX PI De Leys R, Zheng J;
XX DR WPI; 2000-402205/35.
XX CC The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain ESS. HIV is the principle aetiological agent

```

CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. ESS is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 33 AA;

Query Match 92.9%; Score 91; DB 3; Length 33;  
 Best Local Similarity 93.8%; Pred. No. 2.8e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
 |||||:|||||  
 DB 15 LNSWGCKGRIVCYTSV 30

#### RESULT 14

AAB12259  
 ID AAB12259 standard; peptide; 35 AA.

XX AAB12259;

DT 12-SEP-2003 (revised)  
 DT 10-NOV-2000 (first entry)

XX Group O HIV-1 gp41 replacement peptide # 1.

XX HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41.

XX Human immunodeficiency virus 1.

XX EF1013766-A2.

XX 28-JUN-2000.

XX 29-NOV-1999; 99EP-00309491.

XX 30-NOV-1998; 98US-0110292P.

XX 08-FEB-1999; 99US-0119138P.

XX 04-NOV-1999; 99US-00433428.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX De Leys R, Zheng J;

XX WPI; 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for

PT detection of antibodies produced in response to human immunodeficiency

PT virus group O antibodies.

XX Claim 1; Page 37; 52pp; English.

XX The present sequence is a group O Human Immunodeficiency Virus Type 1  
 CC (HIV-1) gp41 peptide. HIV is the principle aetiological agent for  
 CC acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein. The present sequence was derived from a sequence homology  
 CC alignment of group O and group M HIV gp41 descending helix sequences. The  
 CC present sequence was used as a group O replacement peptide, where it was  
 CC used to construct mosaic gp41 proteins, in which the group M  
 CC immunodominant region was replaced by the present sequence. The mosaic  
 CC gp41 proteins (AAB12261 and AAB12262) would be useful as antigens, which  
 CC would be used for the detection of anti-group O HIV antibodies produced  
 CC in response to HIV infection. (Updated on 12-SEP-2003 to standardise OS  
 CC field)

XX SQ Sequence 35 AA;

Query Match 92.9%; Score 91; DB 3; Length 35;

Best Local Similarity 87.5%; Pred. No. 3e-06;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTS 16  
 |||||:|||||  
 DB 16 RLNSWGCKGRICVTS 31

#### RESULT 15

AAW07343

ID AAW07343 standard; peptide; 40 AA.

XX AAW07343;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (ESS).

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.

XX Human immunodeficiency virus 1.

OS WO9627013-A1.

PN 06-SEP-1996.

XX 26-FEB-1996; 96WO-FR000294.

XX 27-FEB-1995; 95FR-00002236.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX N-PSDB; AAT44918.

XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.

XX Claim 12; Page 33; 71pp; French.

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains AN70 and MVPS180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC (NCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF02 (ESS) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)

XX SQ Sequence 40 AA;

Query Match 92.9%; Score 91; DB 2; Length 40;

Best Local Similarity 93.8%; Pred. No. 3.4e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17

Db           |||||:|||||  
          20 LNSWGCKGRIVCYTSV 35

Search completed: May 7, 2004, 18:31:56  
Job time : 37.2836 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:29:27 ; Search time 10.6567 Seconds  
(without alignments)  
82.356 Million cell updates/sec

Title: US-09-147-362A-16

Perfect score: 98

Sequence: 1 RLNSWCKGRLVCYTSV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	94.9	33	3	US-09-433-428D-6
2	93	94.9	33	3	US-09-433-428D-25
3	93	94.9	40	3	US-08-894-699-39
4	93	94.9	40	3	US-08-894-699-68
5	93	94.9	40	3	US-09-444-410-39
6	93	94.9	40	3	US-09-444-410-68
7	91	92.9	23	3	US-09-433-428D-59
8	91	92.9	23	3	US-09-433-428D-60
9	91	92.9	23	3	US-09-433-428D-61
10	91	92.9	28	3	US-09-433-428D-62
11	91	92.9	30	3	US-09-433-428D-63
12	91	92.9	30	3	US-09-433-428D-69
13	91	92.9	33	3	US-09-433-428D-30
14	91	92.9	35	3	US-09-433-428D-64
15	91	92.9	40	3	US-08-894-699-36
16	91	92.9	149	3	US-09-444-410-36
17	91	92.9	149	3	US-09-433-428D-67
18	91	92.9	220	3	US-09-433-428D-66
19	91	92.9	368	3	US-09-433-428D-58
20	91	92.9	439	3	US-09-433-428D-57
21	88	89.8	23	2	US-08-394-021-2
22	88	89.8	23	2	US-08-833-546-4
23	88	89.8	23	3	US-09-388-664-4
24	88	89.8	23	4	US-09-131-551-2
25	88	89.8	23	4	US-09-388-847-4
26	88	89.8	32	4	US-09-625-972-32
27	88	89.8	33	3	US-09-433-428D-2

```

28      88      89.8      35      1      US-08-470-202-62      Sequence 62, Appl
29      88      89.8      35      1      US-08-471-770-62      Sequence 62, Appl
30      88      89.8      35      2      US-08-468-059-62      Sequence 62, Appl
31      88      89.8      35      3      US-09-109-916-62      Sequence 62, Appl
32      88      89.8      35      4      US-09-886-156-62      Sequence 62, Appl
33      88      89.8      35      4      US-09-886-149-62      Sequence 62, Appl
34      88      89.8      35      4      US-09-886-150-62      Sequence 62, Appl
35      88      89.8      35      4      US-09-886-159-62      Sequence 88, Appl
36      88      89.8      37      4      US-08-817-441-88      Sequence 10, Appl
37      88      89.8      146      2      US-08-394-021-10      Sequence 10, Appl
38      88      89.8      146      4      US-09-131-551-10      Sequence 105, App
39      88      89.8      204      3      US-08-965-056-105      Sequence 105, App
40      88      89.8      351      1      US-08-470-202-46      Sequence 46, Appl
41      88      89.8      351      1      US-08-471-770-46      Sequence 46, Appl
42      88      89.8      351      2      US-08-468-059-46      Sequence 46, Appl
43      88      89.8      351      3      US-09-109-916-46      Sequence 46, Appl
44      88      89.8      351      4      US-09-886-156-46      Sequence 46, Appl
45      88      89.8      351      4      US-09-886-149-46      Sequence 46, Appl

```

#### ALIGNMENTS

```

RESULT 1
US-09-433-428D-6
; Sequence 6, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-6

```

```

Query Match      94.9%; Score 93; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2      LNSWCKGRLVCYTSV 17
Db      15      LNSWCKGRLVCYTSV 30

```

```

RESULT 2
US-09-433-428D-25
; Sequence 25, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-25

```

```

Query Match      94.9%; Score 93; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 2 LNSWGCKGRLVCYTSV 17  
| | | | | | | | | | | | | | | | | |  
Db 15 LNSWGCKGRLVCYTSV 30

### RESULT 3

US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCELLEND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

```

,
,
,
,
COMPUTER READABLE FORM:
,
MEDIUM TYPE: Floppy disk
,
COMPUTER: IBM PC compatible
,
OPERATING SYSTEM: PC-DOS/MS-DOS
,
SOFTWARE: Patent In Release #1.0, Version #1.30
,
CURRENT APPLICATION DATA:
,
APPLICATION NUMBER: US/08/894,699
,
FILING DATE: 01-DEC-1997
,
CLASSIFICATION:
,
PRIOR APPLICATION DATA:
,
APPLICATION NUMBER: PCT/FR96/00294
,
FILING DATE: 26-FEB-1996
,

```

PRIOR APPLICATION DATA: 0000  
 APPLICATION NUMBER: FR 95/02236  
 FILING DATE: 27-FEB-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OBLON, NORMAN F.  
 REGISTRATION NUMBER: 24,614  
 REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-413-3000  
 TELEFAX: 703-413-2220  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 40 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-894-699-39

Query Match 94.9%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 3.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels

QY 2 LNSWGCKGRLVCYTSV 17  
| | | | | | | | | |  
Db 20 LNSWGCKGRLVCYTSV 35

## RESULT 4

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
 ; APPLICANT: SARAGOSTI, SENTOB  
 ; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
 ; APPLICANT: LY, THOI-DUONG  
 ; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
 ; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
 ; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 81  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
 ; STREET: FLOOR  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/894,699  
 ; FILING DATE: 01-DEC-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/FR96/00294  
 ; FILING DATE: 26-FEB-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 95/02236  
 ; FILING DATE: 27-FEB-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,614  
 ; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 68:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 40 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-894-699-68

```
Query Match          94.9%; Score 93; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 16: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 2 LNSWGCKGRLVCYTSV 17  
Db 20 LNSWGCKGRLVCYTSV 35

## RESULT 5

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.



STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 94.9%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 3.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
DB 20 LNSWGCKGRLVCYTSV 35

RESULT 6  
US-09-444-410-68  
Sequence 68, Application US/09444410  
Patent No. 6270975  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOI-DUONG  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 94.9%; Score 93; DB 3; Length 40;  
Best Local Similarity 100.0%; Pred. No. 3.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
DB 20 LNSWGCKGRLVCYTSV 35

RESULT 7  
US-09-433-428D-59  
Sequence 59, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
APPLICANT: Zheng, Jian  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433,428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 59  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-59

Query Match 92.9%; Score 91; DB 3; Length 23;  
Best Local Similarity 87.5%; Pred. No. 3.7e-07;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTS 16  
DB 4 RLNSWGCKGRLVCYTS 19

RESULT 8  
US-09-433-428D-60  
Sequence 60, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
APPLICANT: Zheng, Jian

```
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 1
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428D-60
```

```
Query Match          92.9%; Score 91; DB 3; Length 23;
Best Local Similarity 87.5%; Pred. No. 3.7e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RLNSWGCKGRRLVCYTS 16
Db 4 RLNSWGCKGRRIICYTS 19
|||||:|||||
```

```
RESULT 9
US-09-433-428D-61
; Sequence 61, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-433-428D-61
```

```
Query Match          92.9%; Score 91; DB 3; Length 23;
Best Local Similarity 87.5%; Pred. No. 3.7e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RLNSWGCKGRRLVCYTS 16
Db 4 RLNSWGCKGRRIICYTS 19
|||||:|||||
```

```
RESULT 10
US-09-433-428D-62
; Sequence 62, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 28
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 6
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428D-62
```

```
Query Match          92.9%; Score 91; DB 3; Length 28;
Best Local Similarity 87.5%; Pred. No. 4.4e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RLNSWGCKGRRLVCYTS 16
Db 9 RLNSWGCKGRRIICYTS 24
|||||:|||||
```

```
RESULT 11
US-09-433-428D-63
; Sequence 63, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 8
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428D-63
```

```
Query Match          92.9%; Score 91; DB 3; Length 30;
Best Local Similarity 87.5%; Pred. No. 4.7e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RLNSWGCKGRRLVCYTS 16
Db 11 RLNSWGCKGRRIICYTS 26
|||||:|||||
```

```
RESULT 12
US-09-433-428D-69
; Sequence 69, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428D-69
```

```
Query Match          92.9%; Score 91; DB 3; Length 30;
```

Best Local Similarity 87.5%; Pred. No. 4.7e-07;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRILVCYTS 16  
|||||:|||||  
Db 11 RLNSWGCKGRILVCYTS 26

RESULT 13  
US-09-433-428D-30  
; Sequence 30, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-30

Query Match 92.9%; Score 91; DB 3; Length 33;  
Best Local Similarity 93.8%; Pred. No. 5.2e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLNSWGCKGRILVCYTSV 17  
|||||:|||||  
Db 15 RLNSWGCKGRILVCYTSV 30

RESULT 14  
US-09-433-428D-64  
; Sequence 64, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 64  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: Xaa is any amino acid  
; LOCATION: 13  
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-64

Query Match 92.9%; Score 91; DB 3; Length 35;  
Best Local Similarity 87.5%; Pred. No. 5.5e-07;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRILVCYTS 16  
|||||:|||||  
Db 16 RLNSWGCKGRILVCYTS 31

RESULT 15  
US-08-894-699-36  
; Sequence 36, Application US/08894699  
; Patent No. 6030769

GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: PCT/FR96/00294  
; APPLICATION NUMBER: 26-FEB-1996  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-36

Query Match 92.9%; Score 91; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 6.3e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLNSWGCKGRILVCYTSV 17  
|||||:|||||  
Db 20 RLNSWGCKGRILVCYTSV 35

Search completed: May 7, 2004, 18:36:07  
Job time : 11.6567 secs

THE NEW YORK PUBLIC LIBRARY  
ASTOR LENOX TILDEN FOUNDATION  
500 5TH AVENUE  
NEW YORK 17, N.Y.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:34:27 ; Search time 27.1493 Seconds  
(without alignments)  
173.803 Million cell updates/sec

Title: US-09-147-362A-16  
Perfect score: 98  
Sequence: 1 RLNSMGCKGRLVCYTSV 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	89.8	23	10	US-09-388-847-4
2	88	89.8	23	13	US-10-000-321-2
3	88	89.8	32	14	US-10-364-360-32
4	88	89.8	35	9	US-09-886-156-62
5	88	89.8	35	9	US-09-886-150-62
6	88	89.8	35	10	US-09-886-149-62
7	88	89.8	35	10	US-09-886-159-62
8	88	89.8	35	14	US-10-326-090-62
9	88	89.8	37	14	US-10-026-741-88
10	88	89.8	146	13	US-10-000-321-10
11	88	89.8	204	9	US-09-854-816-105
12	88	89.8	351	9	US-09-886-156-46
13	88	89.8	351	9	US-09-886-150-46
14	88	89.8	351	10	US-09-886-149-46
15	88	89.8	351	10	US-09-886-159-46

16	88	89.8	351	14	US-10-326-090-46	Sequence 46, Appl
17	88	89.8	875	14	US-10-369-294-14	Sequence 137, Appl
18	87	88.8	23	14	US-10-320-786-137	Sequence 14, Appl
19	87	88.8	24	14	US-10-320-786-92	Sequence 92, Appl
20	87	88.8	24	14	US-10-320-786-138	Sequence 138, Appl
21	87	88.8	116	14	US-10-320-786-20	Sequence 20, Appl
22	87	88.8	117	14	US-10-320-786-6	Sequence 6, Appl
23	87	88.8	715	14	US-10-320-786-134	Sequence 134, Appl
24	86	87.8	23	14	US-10-320-786-93	Sequence 93, Appl
25	86	87.8	23	14	US-10-320-786-95	Sequence 95, Appl
26	86	87.8	23	14	US-10-320-786-96	Sequence 96, Appl
27	86	87.8	23	14	US-10-320-786-99	Sequence 99, Appl
28	86	87.8	111	14	US-10-320-786-32	Sequence 32, Appl
29	86	87.8	111	14	US-10-320-786-34	Sequence 34, Appl
30	86	87.8	113	14	US-10-320-786-28	Sequence 28, Appl
31	86	87.8	113	14	US-10-320-786-40	Sequence 40, Appl
32	86	87.8	115	14	US-10-320-786-24	Sequence 24, Appl
33	86	87.8	215	8	US-08-911-824-58	Sequence 58, Appl
34	86	87.8	245	8	US-08-911-824-48	Sequence 48, Appl
35	86	87.8	281	8	US-08-911-824-120	Sequence 120, Appl
36	86	87.8	373	8	US-08-911-824-52	Sequence 52, Appl
37	86	87.8	460	8	US-08-911-824-60	Sequence 60, Appl
38	86	87.8	488	8	US-08-911-824-95	Sequence 95, Appl
39	86	87.8	490	8	US-08-911-824-50	Sequence 50, Appl
40	86	87.8	526	8	US-08-911-824-97	Sequence 97, Appl
41	86	87.8	618	8	US-08-911-824-54	Sequence 54, Appl
42	86	87.8	706	8	US-08-911-824-93	Sequence 93, Appl
43	86	87.8	736	8	US-08-911-824-91	Sequence 91, Appl
44	86	87.8	873	8	US-08-911-824-61	Sequence 61, Appl
45	85	86.7	23	14	US-10-320-786-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1  
US-09-388-847-4  
; Sequence 4, Application US/09388847  
; Publication No. US20030004320A1  
; GENERAL INFORMATION:  
; APPLICANT: Amunziato, Michael E  
; Palumbo, Paul S  
; TITLE OF INVENTION: Activated Peptides and Conjugates  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Hoechst Celanese Corporation  
; STREET: Route 202-206, P. O. Box 2500  
; CITY: Somerville  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 08876-1258  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Dell 4100/MXV (IBM PC compatible)  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect for Windows, Version #6.0a  
; CURRENT APPLICATION NUMBER: US/09/388,847  
; FILING DATE: 02-Sep-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,546  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rosenstock, Jerome  
; REGISTRATION NUMBER: 25,456  
; REFERENCE/DOCKET NUMBER: BD11005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 231-2125  
; TELEFAX: (908) 231-4919  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-388-847-4
      89.8%; Score 88; DB 10; Length 23;
Query Match      82.4%; Pred. No. 5.8e-06;
Best Local Similarity
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRIVCYTSV 17
Db 4 RLNLWGCKGLICYTSV 20

RESULT 2
US-10-000-321-2
; Sequence 2, Application US/10000321
; Publication No. US20020123039A1
; GENERAL INFORMATION:
; APPLICANT: BRUST, Stefan
; KNAPP, Stefan
; GERKEN, Manfred
; GUERTLER, Lutz
; TITLE OF INVENTION: Peptides derived from a retrovirus of
; the HIV group, and their use
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,321
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/131,551
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 58315/106/BEAK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-000-321-2
      89.8%; Score 88; DB 13; Length 23;
Query Match      82.4%; Pred. No. 5.8e-06;
Best Local Similarity
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRIVCYTSV 17
Db 4 RLNLWGCKGLICYTSV 20

RESULT 3
```

```
US-10-364-360-32
; Sequence 32, Application US/10364360
; Publication No. US20030180324A1
; GENERAL INFORMATION:
; APPLICANT: GUERTLER, Lutz Gerhard
; APPLICANT: HAUSER, Hans Peter
; APPLICANT: DONGMO DELOKO, Yvette Beatrice
; APPLICANT: ZEKENG, Leopold
; APPLICANT: KAPTUE, Lazare
; TITLE OF INVENTION: LENTIVIRUS FROM THE GROUP OF IMMUNODEFICIENCY VIRUSES OF DRILL MO
; TITLE OF INVENTION: (MANDRILLUS LEUCOPHAERUS) AND THEIR USE
; FILE REFERENCE: 067595/0106
; CURRENT APPLICATION NUMBER: US/10/364,360
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US/09/625,972
; PRIOR FILING DATE: 2000-07-29
; PRIOR APPLICATION NUMBER: DE 199 36 003.0
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 32
; LENGTH: 32
; TYPE: PRT
; ORGANISM: HIV1-O, MVP5180
US-10-364-360-32
      89.8%; Score 88; DB 14; Length 32;
Query Match      82.4%; Pred. No. 7.8e-06;
Best Local Similarity
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRIVCYTSV 17
Db 14 RLNLWGCKGLICYTSV 30

RESULT 4
US-09-886-156-62
; Sequence 62, Application US/09886156
; Patent No. US20020155428A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,156
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-156-62
      89.8%; Score 88; DB 9; Length 35;
Query Match      82.4%; Pred. No. 8.4e-06;
Best Local Similarity
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRIVCYTSV 17
Db 14 RLNLWGCKGLICYTSV 30
```

```
Db      14 RLNLWGCKGKLICTSV 30

; ORGANISM: Human immunodeficiency virus
; US-09-886-149-62

Query Match      89.8%; Score 88; DB 10; Length 35;
Best Local Similarity 82.4%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLNSWGCKGRLVCYTSV 17
      ||| |||||:|||||
Db      14 RLNLWGCKGKLICTSV 30

RESULT 7
US-09-886-159-62
; Sequence 62, Application US/09886159
; Publication No. US20030003443A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,159
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; US-09-886-159-62

Query Match      89.8%; Score 88; DB 10; Length 35;
Best Local Similarity 82.4%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLNSWGCKGRLVCYTSV 17
      ||| |||||:|||||
Db      14 RLNLWGCKGKLICTSV 30

RESULT 8
US-10-326-090-62
; Sequence 62, Application US/10326090
; Publication No. US20030166915A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/10/326,090
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
```

```
Db      14 RLNLWGCKGKLICTSV 30

; ORGANISM: Human immunodeficiency virus
; US-09-886-150-62

Query Match      89.8%; Score 88; DB 9; Length 35;
Best Local Similarity 82.4%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLNSWGCKGRLVCYTSV 17
      ||| |||||:|||||
Db      14 RLNLWGCKGKLICTSV 30

RESULT 6
US-09-886-149-62
; Sequence 62, Application US/09886149
; Publication No. US20030003442A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,149
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; US-09-886-150-62

Query Match      89.8%; Score 88; DB 9; Length 35;
Best Local Similarity 82.4%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLNSWGCKGRLVCYTSV 17
      ||| |||||:|||||
Db      14 RLNLWGCKGKLICTSV 30

RESULT 5
US-09-886-150-62
; Sequence 62, Application US/09886150
; Patent No. US20020172939A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,150
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
; US-09-886-150-62
```

```
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRP
; ORGANISM: Human immunodeficiency virus
US-10-326-090-62

Query Match      89.8%; Score 88; DB 14; Length 35;
Best Local Similarity 82.4%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17
Db 14 RLNLWGCKGKLCYTSV 30

RESULT 9
US-10-026-741-88
; Sequence 88, Application US/10026741
; Publication No. US20030049604A1
; GENERAL INFORMATION:
; APPLICANT: CHARNEAU, PIERRE
; CLAVEL, FRANCOISE
; BORMAN, ANDREW
; QUILLENT, CAROLINE
; GUETARD, DENISE
; MONTAGNIER, LUC
; DONJON DE SAINT-MARTIN, JACQUELINE
; COHEN, JACQUES
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
; SUBTYPE) ANTIGENS
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/026,741
; FILING DATE: 27-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/817,441
; FILING DATE: 31-AUG-1998
; APPLICATION NUMBER: PCT/FR 95/01391
; FILING DATE: 20-OCT-1995
; APPLICATION NUMBER: FR 9412554
; FILING DATE: 20-OCT-1994
; APPLICATION NUMBER: FR 9502526
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03260.6005-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; Molecule Type: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 88:
US-10-026-741-88

Query Match      89.8%; Score 88; DB 14; Length 37;
Best Local Similarity 82.4%; Pred. No. 8.8e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17
Db 14 RLNLWGCKGKLCYTSV 30

RESULT 10
US-10-000-321-10
; Sequence 10, Application US/10000321
; Publication No. US20020123039A1
; GENERAL INFORMATION:
; APPLICANT: BRUST, Stefan
; KNAPPE, Stefan
; GERKEN, Manfred
; GUERTLER, Lutz
; TITLE OF INVENTION: Peptides derived from a retrovirus of
; the HIV group, and their use
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,321
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/131,551
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 58315/106/BEAK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-000-321-10

Query Match      89.8%; Score 88; DB 13; Length 146;
Best Local Similarity 82.4%; Pred. No. 3e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17
Db 60 RLNLWGCKGKLCYTSV 76

RESULT 11
```



US-09-854-816-105  
; Sequence 105, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell  
; J. Christopher Phelan  
; Melissa A. Starovasnik  
; James A. Wells  
; TITLE OF INVENTION: Constrained Helical peptides and Methods of  
; Making Same  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/854,816  
; FILING DATE: 15-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/965,056  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, PhD., Timothy E.  
; REGISTRATION NUMBER: 35,700  
; REFERENCE/DOCKET NUMBER: P1005R2  
; TELEPHONE: 650/225-8674  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 204 amino acids  
; TYPE: Amino acid  
; TOPOLOGY: Linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 105:  
US-09-854-816-105  
  
Query Match 89.8%; Score 88; DB 9; Length 204;  
Best Local Similarity 82.4%; Pred. No. 4.1e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RLNSWGCKGRLVCYTSV 17  
||| |||||:||||  
Db 127 RLNLWGCKGLICYTSV 143  
  
RESULT 12  
US-09-886-156-46  
; Sequence 46, Application US/09886156  
; Patent No. US20020155428A1  
; GENERAL INFORMATION:  
; APPLICANT: Guertler, Lutz G.  
; APPLICANT: Eberle, Josef  
; APPLICANT: Brunn, Albrecht V.  
; APPLICANT: Knapp, Stefan  
; APPLICANT: Hauser, Hans-Peter  
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE  
; FILE REFERENCE: 05495.0001-04  
; CURRENT APPLICATION NUMBER: US/09/886,156  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US/09/109,916  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5  
; PRIOR FILING DATE: 1992-10-06  
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7  
; PRIOR FILING DATE: 1992-10-22  
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8  
; PRIOR FILING DATE: 1992-12-30  
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 46  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-886-156-46

; PRIOR FILING DATE: 1992-10-06  
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7  
; PRIOR FILING DATE: 1992-10-22  
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8  
; PRIOR FILING DATE: 1992-12-30  
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 46  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-886-156-46  
  
Query Match 89.8%; Score 88; DB 9; Length 351;  
Best Local Similarity 82.4%; Pred. No. 6.7e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RLNSWGCKGRLVCYTSV 17  
||| |||||:||||  
Db 97 RLNLWGCKGLICYTSV 113  
  
RESULT 13  
US-09-886-150-46  
; Sequence 46, Application US/09886150  
; Patent No. US20020172939A1  
; GENERAL INFORMATION:  
; APPLICANT: Guertler, Lutz G.  
; APPLICANT: Eberle, Josef  
; APPLICANT: Brunn, Albrecht V.  
; APPLICANT: Knapp, Stefan  
; APPLICANT: Hauser, Hans-Peter  
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE  
; FILE REFERENCE: 05495.0001-04  
; CURRENT APPLICATION NUMBER: US/09/886,150  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/109,916  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5  
; PRIOR FILING DATE: 1992-10-06  
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7  
; PRIOR FILING DATE: 1992-10-22  
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8  
; PRIOR FILING DATE: 1992-12-30  
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 46  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-886-150-46  
  
Query Match 89.8%; Score 88; DB 9; Length 351;  
Best Local Similarity 82.4%; Pred. No. 6.7e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RLNSWGCKGRLVCYTSV 17  
||| |||||:||||  
Db 97 RLNLWGCKGLICYTSV 113  
  
RESULT 14  
US-09-886-149-46  
; Sequence 46, Application US/09886149  
; Publication No. US20030003442A1  
; GENERAL INFORMATION:  
; APPLICANT: Guertler, Lutz G.  
; APPLICANT: Eberle, Josef  
; APPLICANT: Brunn, Albrecht V.

APPLICANT: Knapp, Stefan  
APPLICANT: Hauser, Hans-Peter  
TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE  
FILE REFERENCE: 05495.0001-04  
CURRENT APPLICATION NUMBER: US/09/886,149  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 09/109,916  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: DE P 42 33 646.5  
PRIOR FILING DATE: 1992-10-06  
PRIOR APPLICATION NUMBER: DE P 42 35 718.7  
PRIOR FILING DATE: 1992-10-22  
PRIOR APPLICATION NUMBER: DE P 42 44 541.8  
PRIOR FILING DATE: 1992-12-30  
PRIOR APPLICATION NUMBER: DE P 43 18 186.4  
PRIOR FILING DATE: 1993-06-01  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 46  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-886-149-46

Query Match 89.8%; Score 88; DB 10; Length 351;  
Best Local Similarity 82.4%; Pred. No. 6.7e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
||| |||||:|||||  
Db 97 RLNLWGCKGLICYTSV 113

RESULT 15  
US-09-886-159-46  
Sequence 46, Application US/09886159  
Publication No. US20030003443A1  
GENERAL INFORMATION:  
APPLICANT: Guerlier, Lutz G.  
APPLICANT: Eberle, Josef  
APPLICANT: Brunn, Albrecht V.  
APPLICANT: Knapp, Stefan  
APPLICANT: Hauser, Hans-Peter  
TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE  
FILE REFERENCE: 05495.0001-04  
CURRENT APPLICATION NUMBER: US/09/886,159  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: US/09/109,916  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: DE P 42 33 646.5  
PRIOR FILING DATE: 1992-10-06  
PRIOR APPLICATION NUMBER: DE P 42 35 718.7  
PRIOR FILING DATE: 1992-10-22  
PRIOR APPLICATION NUMBER: DE P 42 44 541.8  
PRIOR FILING DATE: 1992-12-30  
PRIOR APPLICATION NUMBER: DE P 43 18 186.4  
PRIOR FILING DATE: 1993-06-01  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 46  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-886-159-46

Query Match 89.8%; Score 88; DB 10; Length 351;  
Best Local Similarity 82.4%; Pred. No. 6.7e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
||| |||||:|||||  
Db 97 RLNLWGCKGLICYTSV 113

Search completed: May 7, 2004, 18:47:36  
Job time : 27.1493 secs

Query Match 80.6%; Score 79; DB 2; Length 863;  
Best Local Similarity 81.2%; Pred. No. 0.00035;

Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|:|||||:|||||

Db 595 LSLWGCKGRLVCYTSV 610

RESULT 3  
S52930  
Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999  
C:Accession: S52930  
R:Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de  
submitted to the EMBL Data Library, January 1995  
A:Description: A novel HIV-1 O strain illustrates the diversity of the O group.  
A:Reference number: S52929  
A:Accession: S52930  
A>Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-104 <COH>  
A:Cross-references: EMBL:X84328; NID:G695526; PIDN:CAA59066.1; PID:G695527  
C:Superfamily: type E retrovirus env polyprotein

Query Match 76.5%; Score 75; DB 2; Length 104;  
Best Local Similarity 68.8%; Pred. No. 0.00024;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|:|||||:|||||

Db 46 LNLWGRGKAICYTSV 61

RESULT 4  
VCLJG2  
env polyprotein precursor - human immunodeficiency virus type 2 (isolate ROD)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 2, HIV-2  
C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
C:Accession: C26262  
R:Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M.  
Nature 326, 662-669, 1987  
A>Title: Genome organization and transactivation of the human immuno-deficiency virus ty  
A:Reference number: A36262; MUID:87173056; PMID:3031510  
A:Contents: proviral DNA  
A:Accession: C26262  
A:Molecule type: DNA  
A:Residues: 1-858 <GUY>  
A:Cross-references: GB:M15390; NID:g1332361; PIDN:AAB00770.1; PID:g325749  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; immunodeficiency; polyprotein; transmembr

Query Match 70.4%; Score 69; DB 1; Length 858;  
Best Local Similarity 70.6%; Pred. No. 0.011;  
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
|:|||||:|||||

Db 591 RLNSWGCAFRQVCHTTV 607

RESULT 5  
VCLJST  
env polyprotein precursor - human immunodeficiency virus type 2 (isolate ST)  
N:Alternate names: coat polyprotein  
N:Contains: surface glycoprotein gp120; transmembrane glycoprotein gp41

C:Species: human immunodeficiency virus type 2, HIV-2  
A>Note: host Homo sapiens (man)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Sep-1994  
C:Accession: H33943  
R:Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;  
J. Virol. 64, 890-901, 1990  
A>Title: Molecular characterization of an attenuated human immunodeficiency virus type 2  
A:Reference number: A33943; MUID:90112662; PMID:2296086  
A:Accession: H33943  
A:Molecule type: genomic RNA  
A:Residues: 1-859 <KUM>  
A:Cross-references: EMBL:M86924  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-859/Product: env polyprotein #status predicted <ENV>  
F:506-859/Product: surface glycoprotein gp120 #status predicted <SGG>  
F:507-523/Region: hydrophobic  
F:507-523/Region: hydrophobic  
F:675-694/Domain: transmembrane #status predicted <TMN>  
F:36,69,78,113,119,131,137,145,160,173,186,200,232,235,242,266,272,283,294,304,359,392,4

Query Match 69.4%; Score 68; DB 1; Length 859;  
Best Local Similarity 64.7%; Pred. No. 0.016;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
|:|||||:|||||

Db 585 QLNSWGCAFRQVCHTTV 601

RESULT 6  
VCLJG4  
env polyprotein - simian immunodeficiency virus (African green monkey isolate)  
C:Species: simian immunodeficiency virus, SIV  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Oct-1997  
C:Accession: G30045  
R:Fukasawa, M.; Miura, T.; Hasegawa, A.; Morikawa, S.; Tsujimoto, H.; Miki, K.; Kitamura  
Nature 333, 457-461, 1988  
A>Title: Sequence of simian immunodeficiency virus from African green monkey, a new memb  
A:Reference number: A30045; MUID:88232906; PMID:3374586  
A:Accession: G30045  
A:Molecule type: DNA  
A:Residues: 1-864 <FUK>  
A:Cross-references: EMBL:X07805  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; immunodeficiency; polyprotein; transmembr

Query Match 68.4%; Score 67; DB 1; Length 864;  
Best Local Similarity 64.7%; Pred. No. 0.023;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
|:|||||:|||||

Db 616 RLNSWGCAWKQVCHTTV 632

RESULT 7  
S30457  
env protein - human immunodeficiency virus type 2 (fragment)  
C:Species: human immunodeficiency virus type 2, HIV-2  
C>Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999  
C:Accession: S30457  
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
Nature 358, 495-499, 1992  
A>Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
A:Reference number: S30448; MUID:92350299; PMID:1641038  
A:Accession: S30457  
A>Status: translation not shown

A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAO>  
A;Cross-references: EMBL:M87130  
C;Genetics: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein

Query Match 67.3%; Score 66; DB 2; Length 151;  
Best Local Similarity 64.7%; Pred. No. 0.0075;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
:||||| |:::|  
Db 36 KLNSWGCAFRQVCHTTV 52

## RESULT 8

S30456  
env protein - human immunodeficiency virus type 2 (fragment)  
C;Species: human immunodeficiency virus type 2, HIV-2  
C;Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999  
C;Accession: S30456  
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
Nature 358, 495-499, 1992  
A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
A;Reference number: S30448; PMID:92350299; PMID:1641038  
A;Accession: S30456  
A;Status: translation not shown  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAO>  
A;Cross-references: EMBL:M87129  
C;Genetics: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein

Query Match 67.3%; Score 66; DB 2; Length 151;  
Best Local Similarity 64.7%; Pred. No. 0.0075;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
:||||| |:::|  
Db 36 KLNSWGCAFRQVCHTTV 52

## RESULT 9

S30455  
env protein - human immunodeficiency virus type 2 (fragment)  
C;Species: human immunodeficiency virus type 2, HIV-2  
C;Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 23-Mar-2001  
C;Accession: S30455; S30461; S30469  
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
Nature 358, 495-499, 1992  
A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
A;Reference number: S30448; PMID:92350299; PMID:1641038  
A;Accession: S30455  
A;Status: translation not shown  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAO>  
A;Cross-references: EMBL:M87123  
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
submitted to the EMBL Data Library, December 1992  
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
A;Reference number: S30460  
A;Accession: S30461  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GA2>  
A;Cross-references: EMBL:M87131  
A;Accession: S30469  
A;Status: preliminary  
A;Molecule type: nucleic acid

A;Residues: 1-151 <GA3>  
A;Cross-references: EMBL:M87122  
C;Genetics: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein

Query Match 67.3%; Score 66; DB 2; Length 151;  
Best Local Similarity 64.7%; Pred. No. 0.0075;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
:||||| |:::|  
Db 36 KLNSWGCAFRQVCHTTV 52

## RESULT 10

S30454  
env protein - human immunodeficiency virus type 2 (fragment)  
C;Species: human immunodeficiency virus type 2, HIV-2  
C;Date: 09-May-1997 #sequence\_revision 09-May-1997 #text\_change 17-Mar-1999  
C;Accession: S30465; S30460; S30462; S30463; S30464; S30466; S30467; S30468; S30470; S30471  
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
submitted to the EMBL Data Library, December 1992  
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
A;Reference number: S30460  
A;Accession: S30465  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAO>  
A;Cross-references: EMBL:M87128  
A;Accession: S30460  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAW>  
A;Cross-references: EMBL:M87120  
A;Experimental source: clone 22ENVB15  
A;Accession: S30462  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAF>  
A;Cross-references: EMBL:M87132  
A;Experimental source: clone 22ENVB4  
A;Accession: S30463  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAA>  
A;Cross-references: EMBL:M87126  
A;Experimental source: clone 22ENVB3  
A;Accession: S30464  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAZ>  
A;Cross-references: EMBL:M87127  
A;Experimental source: clone 22ENVB4  
A;Accession: S30466  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAY>  
A;Cross-references: EMBL:M87121  
A;Experimental source: clone 22ENVB17  
A;Accession: S30467  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAH>  
A;Cross-references: EMBL:M87119  
A;Experimental source: clone 22ENVB11  
A;Accession: S30468  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAS>  
A;Cross-references: EMBL:M87133  
A;Experimental source: clone 22ENVB5  
A;Accession: S30470  
A;Molecule type: nucleic acid  
A;Residues: 1-151 <GAN>  
A;Cross-references: EMBL:M87134  
A;Experimental source: clone 22ENVB8  
A;Accession: S30471  
A;Molecule type: nucleic acid

A;Residues: 1-151 <GAG>  
 A;Cross-references: EMBL:M87124  
 A;Experimental source: clone 22ENVBB1  
 A;Accession: S30472  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GAG>  
 A;Cross-references: EMBL:M87125  
 A;Experimental source: clone 22ENVBB2  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 Nature 358, 495-499, 1992  
 A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
 A;Reference number: S30448; MUID:92350299; PMID:1641038  
 A;Accession: S30454  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GA2>  
 A;Cross-references: EMBL:M87118  
 A;Superfamily: type E retrovirus env polyprotein

Query Match 67.3%; Score 66; DB 2; Length 151;  
 Best Local Similarity 64.7%; Pred. No. 0.0075; 3; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
 :||||| |::|  
 Db 36 QLNSWGCAFRQVCHTTV 52

RESULT 11  
 S30458  
 env protein - human immunodeficiency virus type 2 (fragment)  
 C;Species: human immunodeficiency virus type 2, HIV-2  
 C;Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 23-Mar-2001  
 C;Accession: S30458; S30477  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 Nature 358, 495-499, 1992  
 A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
 A;Reference number: S30448; MUID:92350299; PMID:1641038  
 A;Accession: S30458  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GAO>  
 A;Cross-references: EMBL:M87142  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 submitted to the EMBL Data Library, December 1992  
 A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
 A;Reference number: S30460  
 A;Accession: S30477  
 A;Status: preliminary  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GA2>  
 A;Cross-references: EMBL:M87141  
 C;Genetics:  
 A;Gene: env  
 C;Superfamily: type E retrovirus env polyprotein  
 C;Keywords: coat protein; glycoprotein

Query Match 66.3%; Score 65; DB 2; Length 151;  
 Best Local Similarity 64.7%; Pred. No. 0.011; 3; Indels 0; Gaps 0;  
 Matches 11; Conservative 3; Mismatches 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
 :||||| |::|  
 Db 36 QLNSWGCAFRQVCHTTV 52

RESULT 12  
 S30459  
 env protein - human immunodeficiency virus type 2 (fragment)  
 C;Species: human immunodeficiency virus type 2, HIV-2  
 C;Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999  
 C;Accession: S30459  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 Nature 358, 495-499, 1992  
 A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.

A;Reference number: S30448; MUID:92350299; PMID:1641038  
 A;Accession: S30459  
 A;Status: translation not shown  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GAO>  
 A;Cross-references: EMBL:M87143  
 C;Genetics:  
 A;Gene: env  
 C;Superfamily: type E retrovirus env polyprotein  
 C;Keywords: coat protein; glycoprotein

Query Match 66.3%; Score 65; DB 2; Length 151;  
 Best Local Similarity 64.7%; Pred. No. 0.011; 3; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
 :||||| |::|  
 Db 36 QLNSWGCAFRQVCHTTV 52

RESULT 13  
 S30448  
 env protein - human immunodeficiency virus type 2 (fragment)  
 C;Species: human immunodeficiency virus type 2, HIV-2  
 C;Date: 02-Dec-1993 #sequence\_revision 30-Jan-1998 #text\_change 17-Mar-1999  
 C;Accession: S30448; S30449; S30480; S30481  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 Nature 358, 495-499, 1992  
 A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.  
 A;Reference number: S30448; MUID:92350299; PMID:1641038  
 A;Accession: S30448  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GAO>  
 A;Cross-references: EMBL:M87069  
 A;Experimental source: FOENVA13  
 A;Accession: S30449  
 A;Status: preliminary; translation not shown  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GAW>  
 A;Cross-references: EMBL:M87071  
 R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 submitted to the EMBL Data Library, December 1992  
 A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
 A;Reference number: S30460  
 A;Accession: S30480  
 A;Status: preliminary  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GA2>  
 A;Cross-references: EMBL:M87085  
 A;Accession: S30481  
 A;Status: preliminary  
 A;Molecule type: nucleic acid  
 A;Residues: 1-151 <GA3>  
 A;Cross-references: EMBL:M87076  
 C;Genetics:  
 A;Gene: env  
 C;Superfamily: type E retrovirus env polyprotein  
 C;Keywords: coat protein; glycoprotein

Query Match 66.3%; Score 65; DB 2; Length 151;  
 Best Local Similarity 64.7%; Pred. No. 0.011; 3; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLVCYTSV 17  
 :||||| |::|  
 Db 36 QLNSWGCAFRQVCHTTV 52

RESULT 14  
 S30453  
 env protein - human immunodeficiency virus type 2 (fragment)  
 C;Species: human immunodeficiency virus type 2, HIV-2

```
C;Date: 02-Dec-1993 #sequence_revision 30-Jan-1998 #text_change 17-Mar-1999
C;Accession: S30453
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
Nature 358, 495-499, 1992
A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.
A;Reference number: S30448; MUID:92350299; PMID:1641038
A;Accession: S30453
A;Status: translation not shown
A;Molecule type: nucleic acid
A;Residues: 1-151 <GAO>
A;Cross-references: EMBL:M87089
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein

Query Match          66.3%; Score 65; DB 2; Length 151;
Best Local Similarity 64.7%; Pred.No. 0.011;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLNSWGCKGRVCYTSV 17
       :||||| |::|:|
Db      36 QLNSWGCAFRQVCHTTV 52

RESULT 15
S30452
env protein - human immunodeficiency virus type 2 (fragment)
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 30-Jan-1998 #text_change 17-Mar-1999
C;Accession: S30452
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
Nature 358, 495-499, 1992
A;Title: Human infection by genetically diverse SIV(SM)-related HIV-2 in West Africa.
A;Reference number: S30448; MUID:92350299; PMID:1641038
A;Accession: S30452
A;Status: translation not shown
A;Molecule type: nucleic acid
A;Residues: 1-151 <GAO>
A;Cross-references: EMBL:M87075
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein

Query Match          66.3%; Score 65; DB 2; Length 151;
Best Local Similarity 64.7%; Pred.No. 0.011;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLNSWGCKGRVCYTSV 17
       :||||| |::|:|
Db      36 QLNSWGCAFRQVCHTTV 52

Search completed: May 7, 2004, 18:35:12
Job time : 8.62687 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:53:36 ; Search time 6.08955 Seconds  
(without alignments)  
145.363 Million cell updates/sec

Title: US-09-147-362A-16  
Perfect score: 98  
Sequence: 1 RLNSWCKGRLVCYTSV 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	75.5	857	1 ENV_HV2KR	Q74126 human immun
2	69	70.4	846	1 ENV_HV2SB	P12449 human immun
3	69	70.4	858	1 ENV_HV2RO	P04577 human immun
4	68	69.4	856	1 ENV_HV2NZ	P05883 human immun
5	67	68.4	854	1 ENV_SIVAT	Q02837 simian immu
6	67	68.4	865	1 ENV_SIVAT	P05896 simian immu
7	65	66.3	712	1 ENV_HV2S2	P32536 human immun
8	65	66.3	851	1 ENV_HV2D1	P17755 human immun
9	65	66.3	851	1 ENV_HV2G1	P18040 human immun
10	65	66.3	854	1 ENV_SIVCZ	P17281 chimpanzee
11	65	66.3	859	1 ENV_HV2CA	P24105 human immun
12	65	66.3	859	1 ENV_HV2D2	P15831 human immun
13	65	66.3	859	1 ENV_HV2ST	P20872 human immun
14	65	66.3	860	1 ENV_HV2BE	P18094 human immun
15	65	66.3	885	1 ENV_SIVS4	P12492 simian immu
16	65	66.3	889	1 ENV_SIVS4	P19503 simian immu
17	64	65.3	768	1 ENV_SIVAG	P27757 simian immu
18	64	65.3	877	1 ENV_SIVAG	P27977 simian immu
19	62	63.3	380	1 ENV_SIVM2	P08810 simian immu
20	62	63.3	881	1 ENV_SIVMK	P05884 simian immu
21	62	63.3	882	1 ENV_SIVM1	P05885 simian immu
22	60	61.2	843	1 ENV_HV1Y2	P35961 human immun
23	60	61.2	847	1 ENV_HV1S1	P19550 human immun
24	60	61.2	847	1 ENV_HV1W2	P05880 human immun
25	60	61.2	851	1 ENV_HV1B8	P04582 human immun
26	60	61.2	852	1 ENV_HV1B8	P12488 human immun
27	60	61.2	852	1 ENV_HV1S3	P19541 human immun
28	60	61.2	853	1 ENV_HV1MF	P19559 human immun
29	60	61.2	853	1 ENV_HV1A2	P12487 human immun
30	60	61.2	855	1 ENV_HV1A2	P03378 human immun
31	60	61.2	855	1 ENV_HV1OY	P20888 human immun
32	60	61.2	855	1 ENV_HV1Z6	P04580 human immun
33	60	61.2	856	1 ENV_HV1B1	P03375 human immun

RESULT 1

ID	ENV_HV2KR	STANDARD;	PRT;	857 AA.
AC	Q74126;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Human immunodeficiency virus type 2 (isolate KR) (HIV-2).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=73484;			
RP	[1]			
RA	SEQUENCE FROM N.A.			
RA	Kraus G.K., Talbot R., Leavitt M., Luzznick L., Schmidt A.,			
RA	Badel P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;			
RL	Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U22047; AAA64582.1; -			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;			
KW	Signal.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	503	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	857	TRANSMEMBRANE GLYCOPROTEIN.
FT	CARBOHYD	36	36	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	69	69	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	78	78	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	113	113	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	121	121	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	142	142	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	159	159	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	198	198	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	264	264	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	281	281	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. .) (POTENTIAL).

ALIGNMENTS

34	60	61.2	856	1	ENV_HV1H2	P04578 human immun
35	60	61.2	856	1	ENV_HV1LW	Q70626 human immun
36	60	61.2	856	1	ENV_HV1MN	P05877 human immun
37	60	61.2	856	1	ENV_HV1PV	P03376 human immun
38	60	61.2	856	1	ENV_HV1SC	P05878 human immun
39	60	61.2	856	1	ENV_HV1W1	P31872 human immun
40	60	61.2	861	1	ENV_HV1BR	P03377 human immun
41	60	61.2	865	1	ENV_HV1RH	P04579 human immun
42	60	61.2	867	1	ENV_HV1J3	P12489 human immun
43	60	61.2	868	1	ENV_HV1C4	P05879 human immun
44	58	59.2	848	1	ENV_HV1JR	P20871 human immun
45	58	59.2	856	1	ENV_HV1H3	P04624 human immun

```
FT CARBOHYD 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 857 AA; 98689 MW; F881C6755B5746DF CRC64;

Query Match 75.5%; Score 74; DB 1; Length 857;
Best Local Similarity 76.5%; Pred. No. 0.00019;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLYCVTSV 17
| | | | | | | | | | | | | | | | | | | |
Db 583 RLNSWGCAFRQVCHTV 599

RESULT 2
ENV HV2SB STANDARD; PRT; 846 AA.
AC P1249,
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate SBLISV) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11718;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89184641; PubMed=2648404;
RA Franchini G., Fagnoli K.A., Glombini F., Jagodzinski L., de Rossi A.,
RA Bosch M., Biberfeld G., Fenyo A.M., Albert J., Gallo R.C.,
RA Wong-Staal F.;
RT Molecular and biological characterization of a replication competent
RT human immunodeficiency type 2 (HIV-2) proviral clone.
RL Proc. Natl. Acad. Sci. U.S.A. 86:2433-2437(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04498; AAB00752.1; -.
DR HIV; J04498; ENV52ISV.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL. 1 19
FT CHAIN 20 493
FT CHAIN 494 846
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 97693 MW; 3A7BE335F914D54C CRC64;

Query Match 70.4%; Score 69; DB 1; Length 846;
Best Local Similarity 70.6%; Pred. No. 0.0013;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRLYCVTSV 17
| | | | | | | | | | | | | | | | | | | |
Db 583 RLNSWGCAFRQVCHTV 599

RESULT 3
ENV HV2RO STANDARD; PRT; 858 AA.
AC P04577;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87173056; PubMed=3031510;
RA Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier L.,
RA Alizon M.;
RT "Genome organization and transactivation of the human
RT immunodeficiency virus type 2.";
RL Nature 326:662-669(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15390; AAB00770.1; -.
DR EMBL; X05291; CAA28914.1; -.
DR PIR; C26262; VCLJG2.
DR HIV; M15390; ENV52ROD.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL. 1 17
FT CHAIN 18 501
FT CHAIN 502 858
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 312 312 T -> I (IN REF. 1; AAB00770).
SQ SEQUENCE 858 AA; 98824 MW; C7266AF1F5C5B9A7 CRC64;

Query Match 70.4%; Score 69; DB 1; Length 858;
Best Local Similarity 70.6%; Pred. No. 0.0013;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGLVCYTSV 17
Db 591 RLNSWGCAFRQVCHTTV 607

RESULT 4
ENV_HV2NZ STANDARD; PRT; 856 AA.
AC P05883;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate NIH-Z) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88320359; PubMed=3261862;
RA Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starcich B.R.,
Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F.,
Arya S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;
RT "Genetic variability between isolates of human immunodeficiency virus
(HIV) type 2 is comparable to the variability among HIV type 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03654; AAB00761.1; -.
DR HIV; J03654; ENV$2NIHZ.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
```

```
FT SIGNAL 1 20
FT CHAIN 21 492 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 493 856 TRANSMEMBRANE GLYCOPROTEIN.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 98665 MW; A938B0A7E2B881D6 CRC64;

Query Match 69.4%; Score 68; DB 1; Length 856;
Best Local Similarity 70.6%; Pred. No. 0.0019;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGLVCYTSV 17
Db 582 RLNSWGCAFRQVCHTTV 598

RESULT 5
ENV_SIVAI STANDARD; PRT; 854 AA.
ID ENV_SIVAI
AC Q02837;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Simian immunodeficiency virus (isolate AGM / clone GRI-1) (SIV-AGM).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=31684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220680; PubMed=2024476;
RA Fomagaard A., Hirsch V.M., Allan J.S., Johnson P.R.;
RT "A highly divergent proviral DNA clone of SIV from a distinct species
of African green monkey.";
RL Virology 182:397-402(1991).
CC -!- MISCELLANEOUS: This is an African green monkey isolate.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M66437; AAA91928.1; -.
DR EMBL; M58410; AAA47591.1; -.

```

DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GPI20.  
 DR Pfam; PF00516; GPI20; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 522  
 FT CHAIN 523 854  
 FT CARBOHYD 35 35  
 FT CARBOHYD 68 68  
 FT CARBOHYD 115 115  
 FT CARBOHYD 136 136  
 FT CARBOHYD 133 133  
 FT CARBOHYD 168 168  
 FT CARBOHYD 182 182  
 FT CARBOHYD 199 199  
 FT CARBOHYD 244 244  
 FT CARBOHYD 255 255  
 FT CARBOHYD 255 255  
 FT CARBOHYD 271 271  
 FT CARBOHYD 283 283  
 FT CARBOHYD 295 295  
 FT CARBOHYD 305 305  
 FT CARBOHYD 355 355  
 FT CARBOHYD 400 400  
 FT CARBOHYD 409 409  
 FT CARBOHYD 458 458  
 FT CARBOHYD 472 472  
 FT CARBOHYD 478 478  
 FT CARBOHYD 623 623  
 FT CARBOHYD 624 624  
 FT CARBOHYD 630 630  
 FT CARBOHYD 646 646  
 SQ SEQUENCE 854 AA; 96855 MW; 5919CA6C9622912F CRC64;

Query Match 68.4%; Score 67; DB 1; Length 854;  
 Best Local Similarity 64.7%; Pred. No. 0.0027;  
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRVLCYTSV 17  
 DB 602 RLNSWGCAKQVCHTV 618  
 RESULT 6  
 ENV\_SIVAT STANDARD; PRT; 865 AA.  
 AC P0586;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GPI20); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Simian immunodeficiency virus (TYO-1 isolate) (SIV-AGM).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11731;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88232906; PubMed=3374586;  
 RA Fukusawa M., Miura T., Hasegawa A., Morikawa S., Tsujimoto H., Miki K., Kitamura T., Hayami M.;  
 RT "Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group.";  
 RL Nature 333:457-461(1988).  
 CC -1- MISCELLANEOUS: This is an African green monkey isolate.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X07805; CAA30663.2; -.  
 DR PIR; G30045; VCLJG4.  
 DR HIV; X07805; ENVAGMTY.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GPI20.  
 DR Pfam; PF00516; GPI20; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 536  
 FT CHAIN 537 865  
 FT SITE 770 770  
 FT CARBOHYD 35 35  
 FT CARBOHYD 68 68  
 FT CARBOHYD 117 117  
 FT CARBOHYD 150 150  
 FT CARBOHYD 165 165  
 FT CARBOHYD 195 195  
 FT CARBOHYD 198 198  
 FT CARBOHYD 210 210  
 FT CARBOHYD 252 252  
 FT CARBOHYD 255 255  
 FT CARBOHYD 266 266  
 FT CARBOHYD 276 276  
 FT CARBOHYD 282 282  
 FT CARBOHYD 294 294  
 FT CARBOHYD 306 306  
 FT CARBOHYD 316 316  
 FT CARBOHYD 373 373  
 FT CARBOHYD 414 414  
 FT CARBOHYD 451 451  
 FT CARBOHYD 488 488  
 FT CARBOHYD 491 491  
 FT CARBOHYD 645 645  
 FT CARBOHYD 661 661  
 SQ SEQUENCE 865 AA; 99024 MW; 6CEFOF09001D6D95 CRC64;

Query Match 68.4%; Score 67; DB 1; Length 865;  
 Best Local Similarity 64.7%; Pred. No. 0.0027;  
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRVLCYTSV 17  
 DB 616 RLNSWGCAKQVCHTV 632  
 RESULT 7  
 ENV\_HV2S2 STANDARD; PRT; 712 AA.  
 AC P32536;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GPI20); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 2 (isolate ST/24.1CH2) (HIV-2).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=31681;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92260681; PubMed=1583738;  
 RA Mulligan M.J., Yarnchikov G.V., Ritter G.D. Jr., Gao F., Jin M.J., Nall C.D., Spies C.P., Hahn B.H., Compans R.W.;  
 RT "Cytoplasmic domain truncation enhances fusion activity by the exterior glycoprotein complex of human immunodeficiency virus type 2 in selected cell types.";  
 RL J. Virol. 66:3971-3975(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M86924; AAA43938.1; --  
 CC PIR; A42535; VCLJ54.  
 CC InterPro; IPR000328; Env GP41.  
 CC InterPro; IPR000777; GP120.  
 CC Pfam; PF00516; GP120; 1.  
 CC Pfam; PF00517; GP41; 1.  
 CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 CC Signal.  
 CC SIGNAL 1 19  
 CC CHAIN 20 712 ENV POLYPROTEIN.  
 CC CHAIN 20 495 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 CC CHAIN 496 712 TRANSMEMBRANE GLYCOPROTEIN.  
 CC TRANSMEM 675 694 POTENTIAL.  
 CC CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 605 605 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 712 AA; 81723 MW; 4EC7F3C83D3C3489 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 712;  
 Best Local Similarity 64.7%; Pred. No. 0.0048;

Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWCGKRLVCYTSV 17  
 :||||| :|||:  
 Db 585 QLNSWCGAFQVCHTTV 601

RESULT 8

ENV\_HV2D1 STANDARD; PRT; 851 AA.  
 AC P17755;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 2 (isolate D194) (HIV-2).

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11713;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91045094; PubMed=22355509;  
 RA Kuehnelt H., Kreutz R., Ruebsamen-Waigmann H.;  
 RT 'Nucleotide sequence of HIV-2D194, an isolate from a Gambian case of  
 RT 'neuro-AIDS', which showed excellent growth in macrophages.';  
 RL Nucleic Acids Res. 18:6142-6142(1990).  
 RN [2]  
 RP SEQUENCE OF 1-266 FROM N.A.  
 RX MEDLINE=89184631; PubMed=2467304;  
 RA Kuehnelt H., von Briesen H., Dietrich U., Adamski M., Mix D.,  
 RA Biersert L., Kreutz R., Immelmann A., Henco K., Meichner C.,  
 RA Andreesen R., Gelderblom H., Ruebsamen-Waigmann H.;  
 RT 'Molecular cloning of two west African human immunodeficiency virus  
 RT type 2 isolates that replicate well in macrophages: a Gambian  
 RT isolate, from a patient with neurologic acquired immunodeficiency  
 RT syndrome, and a highly divergent Ghanaian isolate.';  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:2383-2387(1989).  
 CC -1- MISCELLANEOUS: THIS ISOLATE IS FROM A GAMBIA CASE OF  
 CC 'NEURO-AIDS'.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; J04542; AAA76847.1; --  
 CC EMBL; X52223; CAA36471.1; --  
 CC PIR; S12159; S12159.  
 CC HIV; J04542; ENV52D194.  
 CC InterPro; IPR000328; Env GP41.  
 CC InterPro; IPR000777; GP120.  
 CC Pfam; PF00516; GP120; 1.  
 CC Pfam; PF00517; GP41; 1.  
 CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 CC Signal.  
 CC SIGNAL 1 20  
 CC CHAIN 21 501 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 CC CHAIN 502 851 TRANSMEMBRANE GLYCOPROTEIN.  
 CC CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 626 626 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 851 AA; 97178 MW; 3B002FCD0B9FF118 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 851;  
Best Local Similarity 64.7%; Pred. No. 0.0057;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRVLCVTSV 17  
:||||| |::|  
Db 581 QLNSWGCAFRQVCHTTV 597

RESULT 9  
ENV\_HV2G1 STANDARD; PRT; 851 AA.  
ID ENV\_HV2G1  
AC P18040;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 2 (isolate Ghana-1) (HIV-2).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11717;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9012350; PubMed=2611042;  
RA Hasegawa A., Tsujimoto H., Maki N., Ishikawa K.I., Miura T.,  
RA Fukasawa M., Miki K., Hayami M.;  
RT "Genomic divergence of HIV-2 from Ghana.";  
RL AIDS Res. Hum. Retroviruses 5:593-604 (1989).  
CC -1- MISCELLANEOUS: READTHROUGH OF TERMINATOR UAG BETWEEN CODONS UUG  
CC FOR 738-LEU AND CCU FOR 739-PRO MAY OCCUR.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M30895; AA43931.1; -.  
DR HIV; M30895; ENV52GH1.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.

FT SIGNAL 1 19  
FT CHAIN 20 492 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 493 851 TRANSMEMBRANE GLYCOPROTEIN.  
FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 117 117 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 132 132 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 229 229 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 280 280 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 362 362 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 454 454 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 780 780 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 851 AA; 97491 MW; E6662954B7240F02 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 851;  
Best Local Similarity 64.7%; Pred. No. 0.0057;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRVLCVTSV 17  
:||||| |::|  
Db 582 QLNSWGCAFRQVCHTTV 598

RESULT 10  
ENV\_SIVCZ STANDARD; PRT; 854 AA.  
ID ENV\_SIVCZ  
AC P17281;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90259077; PubMed=2188136;  
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;  
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
RL Nature 345:356-359 (1990).  
CC -1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X52154; CAA36407.1; -.  
DR PIR; S09990; VCLJ51.  
DR HIV; X52154; ENV5CPZ.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
FT SIGNAL 1 30  
FT CHAIN 31 500 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 501 854 TRANSMEMBRANE GLYCOPROTEIN.  
FT TRANSMEM 501 517 POTENTIAL.  
FT TRANSMEM 675 693 POTENTIAL.  
FT TRANSMEM 805 821 POTENTIAL.  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 854;
Best Local Similarity 62.3%; Pred. No. 0.0058;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LNSWGCKGRVLCYTSV 17
   :|||:|:|:|:|:|
Db 583 LGLWGCSGRAVCTTV 598

RESULT 11
ENV HV2CA
ID ENV HV2CA STANDARD; PRT; 859 AA.
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate CAM2) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11715;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91170959; PubMed=2005437;
RA Tristram M., Hill F., Karpas A.;
RT "Nucleotide sequence of a Guinea-Bissau-derived human
RT immunodeficiency virus type 2 proviral clone (HIV-2CAM2).";
RL J. Gen. Virol. 72:721-724(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00835; BAA00716.1; -.
CC PIR; F38475; VCLJCT.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
CC Signal.
CC CHAIN 1 20
CC CHAIN 21 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
CC CHAIN 503 859 TRANSMEMBRANE GLYCOPROTEIN.
CC CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 859 AA; 99018 MW; 6F54913F206B26C3 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 859;
Best Local Similarity 64.7%; Pred. No. 0.0058;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RLNSWGCKGRVLCYTSV 17
   :|||:|:|:|:|
Db 592 QLNSWGCAFRQVCHTTV 608

RESULT 12
ENV HV2D2
ID ENV HV2D2 STANDARD; PRT; 859 AA.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 2 (isolate D205,7) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11716;
RN [1]
RP SEQUENCE FROM N.A.
RA Dietrich U.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-207 FROM N.A.
RX MEDLINE=90081881; PubMed=2594088;
RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehn H.,
RA Ruebsaamen-Waigmann H.;
RT "A highly divergent HIV-2-related isolate.";
RL Nature 342:948-950(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61240; CAA43572.1; -.
CC PIR; S08442; S08442.
CC PIR; S24571; S24571.
CC HIV; X16109; ENV52D205.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
CC Signal.
CC CHAIN 1 23
CC CHAIN 23 23 POTENTIAL.
```



FT CHAIN 24 503 EXTERIOR MEMBRANE GLYCOPROTEIN (BY  
SIMILARITY).  
FT CHAIN 504 859 TRANSMEMBRANE GLYCOPROTEIN  
(BY SIMILARITY).  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 859 AA; 98606 MW; 358C963848733E10 CRC64;

Query Match 66.3%; Score 65; DB 1; Length 859;  
Best Local Similarity 64.7%; Pred. No. 0.0058;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 RLNSWGCKGRLYCYTSV 17  
:||||| |::|:  
Db 593 QLNSWGCAFRQVCHTV 609

RESULT 13  
ENV\_HV2ST  
ID ENV\_HV2ST STANDARD; PRT; 859 AA.  
AC P20872;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 2 (isolate ST) (HIV-2).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11721;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90112862; PubMed=2296086;  
RA Kumar P., Hui H., Kappes J.C., Haggarty B.S., Hoxie J.A., Arya S.K.,  
RA Shaw G.M., Hahn B.H.;  
RT "Molecular characterization of an attenuated human immunodeficiency  
virus type 2 isolate";  
RL J. Virol. 64:890-901(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M31113; AAB01358.1; -.  
DR HIV; M31113; ENV52ST.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 495 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 496 859 TRANSMEMBRANE GLYCOPROTEIN.  
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 859 AA; 99070 MW; B377E79B59A525F CRC64;

Query Match 66.3%; Score 65; DB 1; Length 859;  
Best Local Similarity 64.7%; Pred. No. 0.0058;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 RLNSWGCKGRLYCYTSV 17  
:||||| |::|:  
Db 595 QLNSWGCAFRQVCHTV 601

RESULT 14  
ENV\_HV2BE  
ID ENV\_HV2BE STANDARD; PRT; 860 AA.  
AC P18094;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 2 (isolate BEN) (HIV-2).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11714;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90281594; PubMed=2353457;  
RA Kirchhoff P., Jentsch K., Bachmann B., Stuke A., Laloux C.,  
RA Lueke W., Stahl-Henning C., Schneider J., Nieseit K., Eigen M.,  
RA Hunsmann G.;  
RT "A novel proviral clone of HIV-2: biological and phylogenetic  
relationship to other primate immunodeficiency viruses.";  
RT



RL Virology 177:305-311(1990).  
CC -1- MISCELLANEOUS: THIS ISOLATE IS FROM A GERMAN AIDS PATIENT (WITH  
CC PREDOMINANTLY NEUROLOGICAL COMPLICATIONS) WHO WAS PROBABLY  
CC INFECTED IN MALI.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M30502; AAB00743.1; --  
DR HIV; M30502; ENV52BEN.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 500 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 501 860 TRANSMEMBRANE GLYCOPROTEIN.  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 619 619 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 860 AA; 98931 MW; C7D24EE10136FEDC CRC64;  
Query Match 66.3%; Score 65; DB 1; Length 860;  
Best Local Similarity 64.7%; Pred. No. 0.0058;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RLNSWGCKGRVLCYTSV 17  
:|||||:|:|:|  
Db 590 QLNSWGCAPFQVCHTTV 606  
RESULT 15  
ENV\_SIVS4  
ID ENV\_SIVS4 STANDARD; PRT; 885 AA.  
AC P12492;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxId=11737;

RN SEQUENCE FROM N.A.  
RP MEDLINE=89262053; PubMed=2786147;  
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,  
RA Johnson P.R.;  
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";  
RL Nature 339:389-392(1989).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X14307; CAA32487.1; --  
DR PIR; S04322; S04322.  
DR HIV; X14307; ENV\$SMH4.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane.  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 148 148 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 245 245 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 631 631 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 640 640 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 656 656 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 885 AA; 101863 MW; 7E0D035410D6D988 CRC64;  
Query Match 66.3%; Score 65; DB 1; Length 885;  
Best Local Similarity 64.7%; Pred. No. 0.006;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RLNSWGCKGRVLCYTSV 17  
:|||||:|:|:|  
Db 611 QLNSWGCAPFQVCHTTV 627  
Search completed: May 7, 2004, 18:32:32  
Job time : 6.42289 secs

James M. Smith

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 18:22:06 ; Search time 25.3731 Seconds  
(without alignments)  
211.397 Million cell updates/sec

Title: US-09-147-362A-16  
Perfect score: 98  
Sequence: 1 RLNSWCKGRLVCYTSV 17  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	94.9	116	15 Q7ZJN9	Q7ZJN9 human immun
2	93	94.9	216	15 Q9IEC5	Q9IEC5 human immun
3	93	94.9	219	15 Q9IEB6	Q9IEB6 human immun
4	93	94.9	890	15 Q9Q7G2	Q9Q7G2 human immun
5	91	92.9	219	15 Q9IEC8	Q9IEC8 human immun
6	90	91.8	130	15 Q9IHU9	Q9IHU9 human immun
7	90	91.8	872	15 Q9Q7H0	Q9Q7H0 human immun
8	90	91.8	882	15 Q9Q7F9	Q9Q7F9 human immun
9	90	91.8	887	15 Q9Q7H6	Q9Q7H6 human immun
10	90	91.8	887	15 Q9Q7G9	Q9Q7G9 human immun
11	89	90.8	135	15 Q9DQL9	Q9DQL9 human immun
12	89	90.8	161	15 Q9IEB5	Q9IEB5 human immun
13	89	90.8	242	15 Q9IEB3	Q9IEB3 human immun
14	89	90.8	524	15 Q9IED3	Q9IED3 human immun
15	88	89.8	876	15 Q9Q670	Q9Q670 human immun
16	87	88.8	115	15 Q7ZJN8	Q7ZJN8 human immun

17	87	88.8	116	15	O40459	human immun
18	87	88.8	118	15	O40451	human immun
19	87	88.8	120	15	Q9IHU2	human immun
20	87	88.8	131	15	Q9WR05	human immun
21	87	88.8	131	15	Q9IHU0	human immun
22	87	88.8	137	15	Q9IHV5	human immun
23	87	88.8	146	15	Q9WRV2	human immun
24	87	88.8	213	15	Q9IEC3	human immun
25	87	88.8	214	15	Q9DIK3	human immun
26	87	88.8	214	15	Q9IE96	human immun
27	87	88.8	216	15	Q9IEA5	human immun
28	87	88.8	342	15	O11942	human immun
29	87	88.8	532	15	Q9IEB9	human immun
30	87	88.8	544	15	Q9IED9	human immun
31	87	88.8	548	15	Q9IED6	human immun
32	87	88.8	867	15	Q8Q7G8	human immun
33	87	88.8	871	15	O57073	human immun
34	87	88.8	871	15	O57074	human immun
35	87	88.8	871	15	Q8Q7I2	human immun
36	87	88.8	872	15	Q900Y5	human immun
37	87	88.8	872	15	O57072	human immun
38	87	88.8	900	15	Q9QN28	human immun
39	86	87.8	112	15	O40454	human immun
40	86	87.8	114	15	O40448	human immun
41	86	87.8	116	15	O11941	human immun
42	86	87.8	116	15	O40458	human immun
43	86	87.8	116	15	O40449	human immun
44	86	87.8	117	15	Q9Q6F2	human immun
45	86	87.8	124	15	Q9IHU7	human immun

ALIGNMENTS

RESULT 1

ID Q7ZJN9 PRELIMINARY; PRT; 116 AA.  
AC Q7ZJN9;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=08692A;  
RA Swanson P.A.; Devare S.G., Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
RT integrase, and env gp41."  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214120; AA061840.1; -;  
DR GO; GO:0019031; C:viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 13865 MW; AFIDFCS7E059061 CRC64;

Query Match 94.9%; Score 93; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 4.9e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWCKGRLVCYTSV 17  
|||||  
Db 33 LNSWCKGRLVCYTSV 48

RESULT 2

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5

AC Q9IEC5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.

OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=BCF07;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;

RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ236394; CAB96243.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.

FT NON\_TER 1  
FT NON\_TER 216 216  
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 94.9%; Score 93; DB 15; Length 216;

Best Local Similarity 100.0%; Pred. No. 8.9e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||

Db 48 LNSWGCKGRLVCYTSV 63  
|||||

RESULT 3

Q9IEB6

ID Q9IEB6 PRELIMINARY; PRT; 219 AA.

AC Q9IEB6;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Gp41 (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=BCF14;

RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;

RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ236404; CAB96252.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.

FT NON\_TER 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match

Best Local Similarity 100.0%; Score 93; DB 15; Length 219;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||

Db 58 LNSWGCKGRLVCYTSV 73  
|||||

RESULT 4

Q9Q7G2

ID Q9Q7G2 PRELIMINARY; PRT; 890 AA.

AC Q9Q7G2;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=97US08692A;

RA MEDLINE=21849375; PubMed=11860674;  
RA Yanaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;

RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
RT Phylogenetic Clusters."

RL AIDS Res. Hum. Retroviruses 18:269-282(2002).

DR EMBL; AF383259; AAL98881.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:viral capsid; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.

DR InterPro; IPR000777; Gp120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

QY 2 LNSWGCKGRLVCYTSV 17  
|||||

Db 621 LNSWGCKGRLVCYTSV 636  
|||||

Query Match

Best Local Similarity 100.0%; Score 93; DB 15; Length 890;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||

Db 621 LNSWGCKGRLVCYTSV 636  
|||||

RESULT 5

Q9IEC8

ID Q9IEC8 PRELIMINARY; PRT; 219 AA.

AC Q9IEC8;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Gp41 (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=BCF02;

RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;

RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ236391; CAB96240.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.

FT NON\_TER 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;

Query Match 92.9%; Score 91; DB 15; Length 219;  
Best Local Similarity 93.8%; Pred. No. 2e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||:-:|||||  
DB 57 LNSWGCKGRIVCYTSV 72

## RESULT 6

Q9IHU9 Q9IHU9 PRELIMINARY; PRT; 130 AA.  
AC Q9IHU9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97CM798;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,  
RA Pleniak D., Schable C., Lal R.B.;  
RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O";  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
DR EMBL; AF229235; AAF71912.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 130 130  
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match 91.8%; Score 90; DB 15; Length 130;  
Best Local Similarity 93.8%; Pred. No. 1.8e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||:-:|||||  
DB 41 LNSWGCKGRIVCYTSV 56

## RESULT 7

Q8Q7H0 Q8Q7H0 PRELIMINARY; PRT; 872 AA.  
AC Q8Q7H0;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98CMA453;  
RX MEDLINE=21849375; PubMed=11860674;  
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
Phylogenetic Clusters";  
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
DR EMBL; AF383251; AAL98873.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 872 AA; 98344 MW; 91D37F60053827E3 CRC64;

Query Match 91.8%; Score 90; DB 15; Length 872;  
Best Local Similarity 93.8%; Pred. No. 1.1e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LNSWGCKGRLVCYTSV 17  
|||||:-:|||||  
DB 603 LNSWGCKGRIVCYTSV 618

## RESULT 8

Q8Q7F9 Q8Q7F9 PRELIMINARY; PRT; 882 AA.  
AC Q8Q7F9;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98CMA100;  
RX MEDLINE=21849375; PubMed=11860674;  
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
Phylogenetic Clusters";  
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
DR EMBL; AF383262; AAL98884.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 882 AA; 99439 MW; 627BCA5D0C09C6D4 CRC64;

Query Match 91.8%; Score 90; DB 15; Length 882;  
Best Local Similarity 93.8%; Pred. No. 1.1e-06;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17  
|||||:-:|||||  
DB 614 LNSWGCKGRIVCYTSV 629

## RESULT 9

Q8Q7H6 Q8Q7H6 PRELIMINARY; PRT; 887 AA.  
AC Q8Q7H6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

```

OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CNA124;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383245; AAL98867.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 887 AA; 99366 MW; E210F1E3F7B2474D CRC64;

Query Match 91.8%; Score 90; DB 15; Length 887;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17
|||:||||:|||||
DB 619 LNSWGCKGRLVCYTSV 634

RESULT 10
Q807G9 PRELIMINARY; PRT; 887 AA.
ID Q807G9;
AC Q807G9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CNA124;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383252; AAL98874.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 887 AA; 99559 MW; 88954F4ED76A314A CRC64;

Query Match 91.8%; Score 90; DB 15; Length 887;
Best Local Similarity 93.8%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17
|||:||||:|||||
DB 619 LNSWGCKGRLVCYTSV 634

RESULT 11
Q9DQL9 PRELIMINARY; PRT; 135 AA.
ID Q9DQL9;
AC Q9DQL9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20584646; PubMed=1153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorrquino A.,
RA Vellido J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).
DR EMBL; AF255939; AAG36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match 90.8%; Score 89; DB 15; Length 135;
Best Local Similarity 87.5%; Pred. No. 2.7e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNSWGCKGRLVCYTSV 17
|||:||||:|||||
DB 49 LNSWGCKGRLVCYTSV 64

RESULT 12
Q9IEB5 PRELIMINARY; PRT; 161 AA.
ID Q9IEB5;
AC Q9IEB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF57;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ216405; CAB96253.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
KW NON_TER 1
FT NON_TER 161
SQ SEQUENCE 161 AA; 19305 MW; 5E3AF197E1FDEE7C CRC64;

Query Match 90.8%; Score 89; DB 15; Length 161;
Best Local Similarity 82.4%; Pred. No. 3.2e-07;

```

```
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 RLNSWGCKGRLVCYTSV 17
    ||| |||||:|||||
Db 39 RLNLWGCKGRIICYTSV 55

RESULT 13
ID Q9IE31 PRELIMINARY; PRT; 242 AA.
AC Q9IE31;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TW, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF112;
RA Roques P., Robertson D., Sousquiere S., Damond F., Maucelere P.,
RA Depienne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243366; CAB96336.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 242
SQ SEQUENCE 242 AA; 27539 MW; DE8A73DF0E8A6FD7 CRC64;

Query Match
Best Local Similarity 87.5%; Score 89; DB 15; Length 242;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNSWGCKGRLVCYTSV 17
    ||| |||||:|||||
Db 65 LNSWGCKGRLICYTSV 80

RESULT 14
ID Q9IED3 PRELIMINARY; PRT; 524 AA.
AC Q9IED3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF57;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
RA Depienne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133074; CAB96235.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
```

```
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 524
SQ SEQUENCE 524 AA; 59109 MW; C3D9FF12207AEB41 CRC64;

Query Match
Best Local Similarity 82.4%; Score 89; DB 15; Length 524;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLNSWGCKGRLVCYTSV 17
    ||| |||||:|||||
Db 363 RLNLWGCKGRIICYTSV 379

RESULT 15
ID Q79670 PRELIMINARY; PRT; 876 AA.
AC Q79670;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein gp120/gp41.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94149848; PubMed=8107219;
RA Gurtler L.G., Hauser P.H., Eberle J., von Brunn A., Knapp S.,
RA Zekeng L., Tsague J.M., Kaptue L.;
RT "A new subtype of human immunodeficiency virus type 1 (MVP-5180) from
RT Cameroon.";
RL J. Virol. 68:1581-1585(1994).
DR EMBL; L20571; AAA44864.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 876 AA; 99245 MW; A92F868E37522BAB CRC64;

Query Match
Best Local Similarity 82.4%; Score 88; DB 15; Length 876;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLNSWGCKGRLVCYTSV 17
    ||| |||||:|||||
Db 604 RLNLWGCKGRLICYTSV 620

Search completed: May 7, 2004, 18:34:24
Job time : 25.3731 secs
```

General M. J. H. H. H. H. H.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 41.8739 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-2

Perfect score: 127  
Sequence: 1 LLSLMGCRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	127	100.0	22	2 AAW80460	Peptide d
2	122	96.1	22	2 AAW80465	Peptide d
3	121	95.3	22	2 AAW80466	Peptide d
4	119	93.7	22	2 AAW80459	Peptide d
5	119	93.7	22	2 AAW80471	Peptide d
6	118	92.9	22	2 AAW80461	Peptide d
7	118	92.9	32	2 AAW80470	Peptide d
8	116	91.3	113	2 AAY05565	HIV-1 gro
9	116	91.3	200	3 AAY77373	HIV-1 gro
10	116	91.3	215	2 AAY09499	HIV-1 gro
11	116	91.3	215	2 AAY06983	Recombina
12	116	91.3	215	3 AAY77374	HIV-1 gro
13	116	91.3	245	2 AAY09493	HIV-1 gro
14	116	91.3	245	2 AAY06977	Recombina
15	116	91.3	245	3 AAY77369	HIV-1 gro
16	116	91.3	281	2 AAY09507	HIV-1 gro
17	116	91.3	373	2 AAY09495	HIV-1 gro
18	116	91.3	373	2 AAY06979	Recombina
19	116	91.3	460	2 AAY09500	HIV-1 gro
20	116	91.3	460	2 AAY06984	Recombina
21	116	91.3	460	3 AAY77375	HIV-1 gro
22	116	91.3	474	3 AAY77371	HIV-1 gro
23	116	91.3	488	2 AAY09504	HIV-1 gro
24	116	91.3	490	2 AAY09494	HIV-1 gro
25	116	91.3	490	2 AAY06978	Recombina

26	116	91.3	490	3 AAY77370	HIV-1 gro
27	116	91.3	526	2 AAY09505	HIV-1 gro
28	116	91.3	618	2 AAY09496	HIV-1 gro
29	116	91.3	618	2 AAY06980	Recombina
30	116	91.3	618	3 AAY77372	HIV-1 gro
31	116	91.3	706	2 AAY09503	HIV-1 gro
32	116	91.3	715	2 AAY05625	HIV-1 gro
33	116	91.3	736	2 AAY09502	HIV-1 gro
34	116	91.3	873	2 AAY09501	HIV-1 gro
35	116	91.3	873	2 AAY06985	Amino aci
36	116	91.3	873	3 AAY77376	HIV-1 gro
37	115	90.6	104	2 AAW07245	HIV-1 gro
38	114	89.8	32	2 AAW80469	Peptide d
39	113	89.0	22	2 AAW80462	Peptide d
40	113	89.0	113	2 AAY05559	HIV-1 gro
41	113	89.0	116	2 AAY05555	HIV-1 gro
42	113	89.0	356	2 AAW03940	GP 41 ant
43	112	88.2	113	2 AAY05546	HIV-1 gro
44	112	88.2	113	2 AAY05554	HIV-1 gro
45	112	88.2	117	2 AAY05548	HIV-1 gro

ALIGNMENTS

RESULT 1  
AAW80460  
ID AAW80460 standard; peptide; 22 AA.  
XX AAW80460;  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
DE Peptide derived from a conserved sequence of group O human HIV.  
XX  
KW Group O human immune deficiency virus; HIV; detection; infection.  
XX  
OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX  
PN WO9845323-A1.  
XX  
PD 15-OCT-1998.  
XX  
PF 06-APR-1998; 98WO-FR000691.  
XX  
PR 09-APR-1997; 97FR-00004356.  
XX 24-FEB-1998; 98FR-00002212.  
XX  
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX  
PI Cheneboux DMB, Delagneau JH, Gadelle SUX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
XX immune deficiency virus of group O.  
XX  
PT Claim 6; Page 42; 55pp; French.  
XX  
PS AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX connected around short highly conserved sequences present in isolates of  
XX group O human immune deficiency virus (HIV). The peptides are useful as  
XX immunological reagents for detecting infection by group O human immune  
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX (Updated on 27-AUG-2003 to correct OS field.)  
SQ Sequence 22 AA;

Query Match 100.0%; Score 127; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 2.8e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQVNET 22  
Db 1 LLSLMGCRGLVCYTSVQVNET 22

RESULT 2  
AAW80465  
ID AAW80465 standard; peptide; 22 AA.  
XX AC AAW80465;  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX Peptide derived from a conserved sequence of group O human HIV.  
XX Group O human immune deficiency virus; HIV; detection; infection.  
XX Synthetic.  
OS Human immunodeficiency virus 1.  
XX WO9845323-A1.  
XX 15-OCT-1998.  
XX 06-APR-1998; 98WO-FR000691.  
XX 09-APR-1997; 97FR-00004356.  
DT 24-FEB-1998; 98FR-00002212.  
DT 28-JAN-1999 (first entry)  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
PI WPI; 1998-583190/49.  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
PT Claim 6; Page 43; 55pp; French.  
PS AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX connected around short highly conserved sequences present in isolates of  
XX group O human immune deficiency virus (HIV). The peptides are useful as  
XX immunological reagents for detecting infection by group O human immune  
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX (Updated on 27-AUG-2003 to correct OS field.)  
XX Sequence 22 AA;

Query Match 96.1%; Score 122; DB 2; Length 22;  
Best Local Similarity 95.5%; Pred. No. 1.3e-09;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQVNET 22  
Db 1 LLSLMGCRGAVCYTSVQVNET 22

RESULT 3  
AAW80466  
ID AAW80466 standard; peptide; 22 AA.  
XX AC AAW80466;  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)

XX Peptide derived from a conserved sequence of group O human HIV.  
DE Group O human immune deficiency virus; HIV; detection; infection.  
XX Synthetic.  
OS Human immunodeficiency virus 1.  
XX WO9845323-A1.  
XX 15-OCT-1998.  
XX 06-APR-1998; 98WO-FR000691.  
XX 09-APR-1997; 97FR-00004356.  
DT 24-FEB-1998; 98FR-00002212.  
DT 28-JAN-1999 (first entry)  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
PI WPI; 1998-583190/49.  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
PT Claim 6; Page 43; 55pp; French.  
PS AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX connected around short highly conserved sequences present in isolates of  
XX group O human immune deficiency virus (HIV). The peptides are useful as  
XX immunological reagents for detecting infection by group O human immune  
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX (Updated on 27-AUG-2003 to correct OS field.)  
XX Sequence 22 AA;

Query Match 95.3%; Score 121; DB 2; Length 22;  
Best Local Similarity 95.5%; Pred. No. 1.8e-09;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQVNET 22  
Db 1 LLSLMGCRGLVCYTSVQVNET 22

RESULT 4  
AAW80459  
ID AAW80459 standard; peptide; 22 AA.  
XX AC AAW80459;  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX Peptide derived from a conserved sequence of group O human HIV.  
DE Group O human immune deficiency virus; HIV; detection; infection.  
XX Synthetic.  
OS Human immunodeficiency virus 1.  
XX WO9845323-A1.  
XX 15-OCT-1998.  
XX 06-APR-1998; 98WO-FR000691.  
XX 09-APR-1997; 97FR-00004356.  
DT 24-FEB-1998; 98FR-00002212.  
DT 28-JAN-1999 (first entry)

```

PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX
DR WPI; 1998-583190/49.
XX
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
XX Claim 6; Page 42; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 22 AA;

Query Match 93.7%; Score 119; DB 2; Length 22;
Best Local Similarity 90.9%; Pred. No. 3.3e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 1 LLSLWGCGRGKAVCYTSVQWNET 22

RESULT 5
AAW80471
ID AAW80471 standard; peptide; 32 AA.
XX
AC AAW80471;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 44; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 22 AA;

Query Match 92.9%; Score 118; DB 2; Length 22;
Best Local Similarity 90.9%; Pred. No. 4.5e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 1 LLSLWGCGRGLVCYTSVQWNET 22

RESULT 7
AAW80470
ID AAW80470 standard; peptide; 32 AA.
XX
XX

```

CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 32 AA;

Query Match 93.7%; Score 119; DB 2; Length 32;  
 Best Local Similarity 90.9%; Pred. No. 4.8e-09;  
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
 DB 11 LLDLWGCGRGLVCYTSVRWNET 32

RESULT 6  
 AAW80461  
 ID AAW80461 standard; peptide; 22 AA.  
 XX  
 AC AAW80461;

XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX

DE Peptide derived from a conserved sequence of group O human HIV.  
 XX  
 KW Group O human immune deficiency virus; HIV; detection; infection.

XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 XX WO9845323-A1.

XX 15-OCT-1998.  
 XX  
 XX 06-APR-1998; 98WO-FR000691.

XX 09-APR-1997; 97FR-00004356.  
 XX 24-FEB-1998; 98FR-00002212.

XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;

XX WPI; 1998-583190/49.

XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.

XX Claim 6; Page 42; 55pp; French.

XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 XX (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 22 AA;

Query Match 92.9%; Score 118; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 4.5e-09;  
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
 DB 1 LLSLWGCGRGLVCYTSVQWNET 22

RESULT 7  
 AAW80470  
 ID AAW80470 standard; peptide; 32 AA.  
 XX

```

AC AAW80470;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX Peptide derived from a conserved sequence of group O human HIV.
DE Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
OS Human immunodeficiency virus 1.
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
XX 24-FEB-1998; 98FR-00002212.
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gabelle SUX, Rieunier FY;
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
XX Claim 6; Page 44; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 32 AA;
SQ
Query Match 92.9%; Score 118; DB 2; Length 32;
Best Local Similarity 86.4%; Pred. No. 6.5e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 3;
QY 1 LLSLWGCRGLVCYTSVQWNET 22
Db 11 LLNINWCRGLVCYTSVRWNET 32
RESULTS
AAW05565
ID AAY05565 standard; protein; 113 AA.
XX
XX AAY05565;
AC
XX
XX 17-OCT-2003 (revised)
DT 19-JUL-1999 (first entry)
XX
XX HIV-1 group O isolate MP539-PBMC gp41 antigen.
DE
XX HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine;
XX diagnosis; AIDS.
KW
XX Human immunodeficiency virus 1.
OS
XX
XX Key Location/Qualifiers
FH
XX Misc-difference 65 /note= "encoded by AMW"
FT
XX Misc-difference 74 /note= "encoded by ATR"
FT
XX

```

---

```

FT Misc-difference 84 /note= "encoded by GAK"
FT
FT Misc-difference 86 /note= "encoded by AGY"
FT
XX
XX WO9904011-A2.
XX
XX 28-JAN-1999.
XX
XX 20-JUL-1998; 98WO-EP004522.
XX
XX 18-JUL-1997; 97EP-00870110.
XX (INNO-) INNOGENETICS NV.
XX
XX Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;
XX WPI; 1999-132255/11.
XX N-PSDB; AAY05565.
XX
XX New isolated HIV-1 group O strains - used to produce polynucleotides,
PT antigens and antibodies for use in diagnosis and in vaccines for
PT prevention of HIV-1 infection.
XX
XX Claim 3; Fig 6; 162pp; English.
XX
XX The present sequence is an antigen of the gp41 protein of HIV-1 group O
CC (Outlier) strain MP539-PBMC, a Cameroon isolate. The invention relates to
CC new HIV-1 group O antigens (see AAY05546-625), and the use of these
CC antigens, or nucleic acids encoding them (see AAX25154-80), in the
CC diagnosis and prophylaxis of AIDS. They can be used as reagents for
CC detecting HIV-1 group O infection and for differentiating different types
CC of HIV-1 group O infection. Vaccines that provide protective immunity
CC against HIV-1 infection, in particular against HIV-1 group O infection,
CC comprise at least one HIV-1 type O antigen, a nucleic acid encoding such
CC an antigen, a virus-like particle comprising such an antigen, or an
CC attenuated form of an HIV-1 type O strain. The invention also relates to
CC new HIV-1 group O strains, mostly from patients from Cameroon and its
CC neighbouring countries. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 113 AA;
SQ
Query Match 91.3%; Score 116; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 4.1e-08;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 LLSLWGCRGLVCYTSVQWNET 22
Db 39 LLNLWGCKGRILCYTSVQWNET 60
RESULTS
AAW77373
ID AAY77373 standard; protein; 200 AA.
XX
XX AAY77373;
AC
XX
XX 22-MAY-2000 (first entry)
DT
XX
XX HIV-1 group O env gp120/gp41 pGO-5 recombinant protein, SEQ ID NO:56.
DE
XX HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;
KW immunoassay; positive control; affinity purification; therapeutic;
KW Escherichia coli; antigen; synthetic gene construction; mutagen;
KW deletion mutation.
XX
XX Human immunodeficiency virus 1; group O isolate HAM112.
OS
XX Synthetic.
XX
XX WO200004383-A2.
XX
XX 27-JAN-2000.
XX

```

PF 09-JUL-1999; 99WO-US015469.  
 XX  
 PR 14-JUL-1998; 98US-00115171.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 XX Scheffel JW, Hackett JR, Tyner JD, Hickman RK;  
 PI WPI; 2000-171290/15.  
 XX N-PSDB; AAZ90284.  
 DR  
 XX Novel monoclonal antibodies useful as positive control reagent for  
 PT detecting human immunodeficiency virus infections and diagnosing,  
 PT evaluating or prognosing viral disease.  
 XX  
 PS Example 3; Fig 9; 148pp; English.  
 XX  
 CC The invention relates to anti-HIV-1 group O monoclonal antibodies, which  
 CC may be used as positive control reagents in immunoassays to detect and  
 CC differentiate HIV-1 infections. The invention also encompasses a  
 CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,  
 CC which has no more than 15% cross reactivity to a corresponding antigen  
 CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of  
 CC using a monoclonal antibody as a positive control reagent in an  
 CC immunoassay for the detection of anti HIV-1 group O antibodies. The  
 CC monoclonal antibodies are useful as positive control reagents in  
 CC immunoassays capable of detecting anti-HIV-1 group O antibodies. Such  
 CC immunoassays involve coupling a monoclonal antibody with HIV group-1  
 CC antigen and detecting the antigen-antibody complex. The monoclonal  
 CC antibodies of the invention would be used to ensure that the reagents  
 CC provided to detect HIV-1 group O antibody were performing properly. The  
 CC monoclonal antibodies may also can be immobilised on a matrix and used  
 CC for affinity purification of specific HIV-1 group O-derived proteins from  
 CC cell cultures or biological tissues. The monoclonal antibodies can also  
 CC be used for generating chimeric antibodies for therapeutic use. Different  
 CC synthetic, recombinant or purified antibodies which identify different  
 CC epitopes of HIV antigens can be used in combination in assay to diagnose,  
 CC evaluate, or prognosticate HIV disease condition. The monoclonal  
 CC antibodies are also useful for differentiating HIV-1 Group O antigens  
 CC from HIV-group M and HIV-2 antigens. Sequences AAY7369-Y7735 represent  
 CC recombinant HIV-1 group O env antigens encoded by the synthetic genes  
 CC AAZ90280-Z90286. The recombinant HIV-1 env proteins contain various  
 CC deletions relative to the native HAM12 isolate env protein (AAY77376).  
 CC The recombinant HIV-1 group O antigens were purified and used to screen  
 CC hybridoma cultures  
 XX  
 SQ Sequence 200 AA;  
 Query Match 91.3%; Score 116; DB 3; Length 200;  
 Best Local Similarity 81.8%; Pred. No. 7e-08;  
 Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSLMGCRGLVCYTSVQWNET 22  
 DB 112 LLLNLGCKGRLLCYTSVKWNET 133  
 RESULT 10  
 AAY09499  
 ID AAY09499 standard; protein; 215 AA.  
 XX  
 AC AAY09499;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE HIV-1 Group O env polypeptide pGO-8PL.  
 XX  
 XX HIV; human immunodeficiency virus; antigen; detection; antibody;  
 KW differentiation; Group O; env; immunogen; immunoassay.  
 XX  
 OS Human immunodeficiency virus 1.  
 FT  
 XX

PN WO9909179-A2.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 17-AUG-1998; 98WO-US017014.  
 XX  
 PR 15-AUG-1997; 97US-00911824.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;  
 XX WPI; 1999-190167/16.  
 DR N-PSDB; AAX56078.  
 XX  
 XX New isolated HIV-1 Group O env polypeptides - used for the detection of  
 PT anti-HIV antibodies and for the production of antibodies for use in  
 PT detection, purification and therapy.  
 PS  
 XX Claim 17; Fig 5; 138pp; English.  
 CC The present invention describes (A) an isolated HIV-1 Group O env  
 CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env  
 CC polypeptide comprising an immunoreactive portion of a polypeptide as in  
 CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);  
 CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide  
 CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct  
 CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at  
 CC least one HIV-1 Group M env polypeptide; (5) an antigen construct  
 CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env  
 CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an  
 CC antigen construct comprising a first HIV-2 env polypeptide fused to a  
 CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as  
 CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a  
 CC host cell transformed by an expression vector as in (8); and (10) an  
 CC immunoassay kit for the detection of antibodies to HIV-1 comprising an  
 CC antigen construct as in (3)-(6). The antigen constructs can be used for  
 CC the detection of anti-HIV-1 antibodies in test samples. They can also be  
 CC used as immunogens to produce antibodies. The antibodies can be used to  
 CC purify HIV polypeptides, for therapy and for detection of HIV  
 CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 215 AA;  
 Query Match 91.3%; Score 116; DB 2; Length 215;  
 Best Local Similarity 81.8%; Pred. No. 7.5e-08;  
 Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSLMGCRGLVCYTSVQWNET 22  
 DB 127 LLLNLGCKGRLLCYTSVKWNET 148  
 RESULT 11  
 AAY06983  
 ID AAY06983 standard; protein; 215 AA.  
 XX  
 AC AAY06983;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Recombinant pGO-8PL protein.  
 XX  
 KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;  
 KW antibody; assay.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 XX Key Location/Qualifiers  
 FT Protein 2..46  
 FT Peptide /note= "gp120 sequence"  
 FT 47..245

```

FT XX /note= "gp41 sequence"
PN XX WO9509410-A2.
XX XX
PD XX 25-FEB-1999.
XX XX
XX PF 07-AUG-1998; 98WO-US016506.
XX XX
XX PR 15-AUG-1997; 97US-00912129.
XX XX
XX PA (ABBO ) ABBOTT LAB.
XX XX
XX PI Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;
XX PI Golden AM, Brennan CA, Devare SG;
XX XX
XX DR WPI; 1999-190224/16.
XX DR N-PSDB; AAX37193.
XX XX
XX PT New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can
XX PT be used in field assay, requiring no electricity and less specialised
XX PT equipment.
XX PS Claim 1; Fig 5; 104pp; English.
XX XX
XX CC The invention relates to a rapid assay for simultaneous detection and
XX CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The
XX CC method comprises (a) contacting the sample with a strip containing at
XX CC least one immobilised capture reagent per analyte and on which the sample
XX CC moves from the proximal to the distal end by capillary action, under
XX CC conditions sufficient to form capture reagent/analyte complexes, and (b)
XX CC determining the presence of analyte(s) by detecting a visible colour
XX CC change at the capture reagent site on the strip wherein the capture
XX CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80
XX CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown
XX CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in
XX CC AAY06981. The invention is used to screen patients for antibodies to HIV-
XX CC 1 types O and M, and HIV-2. The invention will be particularly useful in
XX CC places and situation where equipment and/or electricity is not available.
XX CC The invention provides a screening method which is faster and requires a
XX CC less equipment than prior art methods. The present sequence represents a
XX CC amino acid sequence of the recombinant pGO-9PL protein which acts as a
XX CC capture reagent for HIV-1 group O
XX SQ Sequence 215 AA;
XX
XX Query Match 91.3%; Score 116; DB 2; Length 215;
XX Best Local Similarity 81.8%; Pred. No. 7.5e-08;
XX Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LLSLWGCRGLVCYTSVQWNET 22
XX ||:||||:||||:||||:||||
XX Db 127 LNLWGCKRGLICYTSVKWNET 148
XX
XX RESULT 12
XX AAY77374
XX ID AAY77374 standard; protein; 215 AA.
XX XX
XX AC AAY77374;
XX XX
XX DT 22-MAY-2000 (first entry)
XX XX
XX DE HIV-1 group O env gp120/gp41 pGO-9PL recombinant protein, SEQ ID NO:58.
XX XX
XX KW HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;
XX KW immunoassay; positive control; affinity purification; therapeutic;
XX KW Escherichia coli; antigen; synthetic gene construction; muten;
XX KW deletion mutation.
XX XX
XX OS Human immunodeficiency virus 1; group O isolate HAM112.
XX OS Synthetic.
XX XX
XX PN WO200004383-A2.

```

---

```

XX PD 27-JAN-2000.
XX XX
XX PF 09-JUL-1999; 99WO-US015469.
XX XX
XX PR 14-JUL-1998; 98US-00115171.
XX XX
XX PA (ABBO ) ABBOTT LAB.
XX XX
XX PI Scheffel JW, Hackett JR, Tyner JD, Hickman RK;
XX XX
XX DR WPI; 2000-171290/15.
XX DR N-PSDB; AAZ90285.
XX XX
XX PT Novel monoclonal antibodies useful as positive control reagent for
XX PT detecting human immunodeficiency virus infections and diagnosing,
XX PT evaluating or prognosing viral disease.
XX XX
XX PS Example 3; Fig 5; 148pp; English.
XX XX
XX CC The invention relates to anti-HIV-1 group O monoclonal antibodies, which
XX CC may be used as positive control reagents in immunoassays to detect and
XX CC differentiate HIV-1 infections. The invention also encompasses a
XX CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,
XX CC which has no more than 15% cross reactivity to a corresponding antigen
XX CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of
XX CC using a monoclonal antibody as a positive control reagent in an
XX CC immunoassay for the detection of anti HIV-1 group O antibodies. The
XX CC monoclonal antibodies are useful as positive control reagents in
XX CC immunoassays capable of detecting anti-HIV-1 Group O antibodies. Such
XX CC immunoassays involve coupling a monoclonal antibody with HIV group-1
XX CC antigen and detecting the antigen-antibody complex. The monoclonal
XX CC antibodies of the invention would be used to ensure that the reagents
XX CC provided to detect HIV-1 group O antibody were performing properly. The
XX CC monoclonal antibodies may also be immobilised on a matrix and used
XX CC for affinity purification of specific HIV-1 group O-derived proteins from
XX CC cell cultures or biological tissues. The monoclonal antibodies can also
XX CC be used for generating chimeric antibodies for therapeutic use. Different
XX CC synthetic, recombinant or purified antibodies which identify different
XX CC epitopes of HIV antigens can be used in combination in assay to diagnose,
XX CC evaluate, or prognosticate HIV disease condition. The monoclonal
XX CC antibodies are also useful for differentiating HIV-1 Group O antigens
XX CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent
XX CC recombinant HIV-1 group O env antigens encoded by the synthetic genes
XX CC AAZ90280-290286. The recombinant HIV-1 env proteins contain various
XX CC deletions relative to the native HAM112 isolate env protein (AAY77376).
XX CC The recombinant HIV-1 group O antigens were purified and used to screen
XX CC hybridoma cultures
XX SQ Sequence 215 AA;
XX
XX Query Match 91.3%; Score 116; DB 3; Length 215;
XX Best Local Similarity 81.8%; Pred. No. 7.5e-08;
XX Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LLSLWGCRGLVCYTSVQWNET 22
XX ||:||||:||||:||||:||||
XX Db 127 LNLWGCKRGLICYTSVKWNET 148
XX
XX RESULT 13
XX ID AAY09493 standard; protein; 245 AA.
XX XX
XX AC AAY09493;
XX XX
XX DT 17-OCT-2003 (revised)
XX DT 15-JUL-1999 (first entry)
XX XX
XX DE HIV-1 Group O env polypeptide pGO-9PL.
XX XX
XX KW HIV; human immunodeficiency virus; antigen; detection; antibody;
XX KW differentiation; Group O; env; immunogen; immunoassay.

```

XX OS Human immunodeficiency virus 1.  
 XX PN WO9909179-A2.  
 XX PD 25-FEB-1999.  
 XX PF 17-AUG-1998; 98WO-US017014.  
 XX PR 15-AUG-1997; 97US-00911824.  
 XX PA (ABBO ) ABBOTT LAB.  
 XX PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;  
 XX DR N-PSDB; AAX56074.  
 XX PT New isolated HIV-1 Group O env polypeptides - used for the detection of  
 XX PT anti-HIV antibodies and for the production of antibodies for use in  
 XX PT detection, purification and therapy.  
 XX PS Claim 16; Fig 7; 138pp; English.  
 XX CC The present invention describes (A) an isolated HIV-1 Group O env  
 XX CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env  
 XX CC polypeptide comprising an immunoreactive portion of a polypeptide as in  
 XX CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);  
 XX CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide  
 XX CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct  
 XX CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at  
 XX CC least one HIV-1 Group M env polypeptide; (5) an antigen construct  
 XX CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env  
 XX CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an  
 XX CC antigen construct comprising a first HIV-2 env polypeptide fused to a  
 XX CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as  
 XX CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a  
 XX CC host cell transformed by an expression vector as in (8); and (10) an  
 XX CC immunoassay kit for the detection of antibodies to HIV-1 comprising an  
 XX CC antigen construct as in (3)-(6). The antigen constructs can be used for  
 XX CC the detection of anti-HIV-1 antibodies in test samples. They can also be  
 XX CC used as immunogens to produce antibodies. The antibodies can be used to  
 XX CC purify HIV polypeptides, for therapy and for detection of HIV  
 XX CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)  
 XX SQ Sequence 245 AA;  
 Query Match 91.3%; Score 116; DB 2; Length 245;  
 Best Local Similarity 81.8%; Pred. No. 8.5e-08;  
 Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
 Db 127 LLNLWCKGRGLICYTSVKWNET 148  
 RESULT 14  
 AAY06977  
 ID AAY06977 standard; protein; 245 AA.  
 XX AC AAY06977;  
 XX XX  
 XX DT 06-JUL-1999 (first entry)  
 XX DE Recombinant pGO-9PL protein.  
 XX HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;  
 KW antibody; assay.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX Human immunodeficiency virus 1; group O isolate HAM112.  
 XX FH Key Location/Qualifiers

FT Protein 2. .46  
 FT FT /note= "gp120 sequence"  
 FT Peptide 47. .245  
 XX /note= "gp41 sequence"  
 XX PN WO9909410-A2.  
 XX PD 25-FEB-1999.  
 XX PF 07-AUG-1998; 98WO-US016506.  
 XX PR 15-AUG-1997; 97US-00912129.  
 XX PA (ABBO ) ABBOTT LAB.  
 XX PI Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;  
 XX DR Golden AM, Brennan CA, Devare SG;  
 XX DR WPI; 1999-190224/16.  
 XX DR N-PSDB; AAX37189.  
 XX PT New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can  
 XX PT be used in field assay, requiring no electricity and less specialised  
 XX PT equipment.  
 XX CC Claim 1; Fig 7; 104pp; English.  
 XX CC The invention relates to a rapid assay for simultaneous detection and  
 XX CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The  
 XX CC method comprises (a) contacting the sample with a strip containing at  
 XX CC least one immobilised capture reagent per analyte and on which the sample  
 XX CC moves from the proximal to the distal end by capillary action, under  
 XX CC conditions sufficient to form capture reagent/analyte complexes, and (b)  
 XX CC determining the presence of analyte(s) by detecting a visible colour  
 XX CC change at the capture reagent site on the strip wherein the capture  
 XX CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06377-80  
 XX CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown  
 XX CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in  
 XX CC AAY06981. The invention is used to screen patients for antibodies to HIV-  
 XX CC 1 types O and M, and HIV-2. The invention will be particularly useful in  
 XX CC places and situation where equipment and/or electricity is not available.  
 XX CC The invention provides a screening method which is faster and requires a  
 XX CC less equipment than prior art methods. The present sequence represents a  
 XX CC amino acid sequence of the recombinant pGO-9PL recombinant protein which  
 XX CC acts as a capture reagent for HIV-1 group O  
 XX SQ Sequence 245 AA;  
 Query Match 91.3%; Score 116; DB 2; Length 245;  
 Best Local Similarity 81.8%; Pred. No. 8.5e-08;  
 Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
 Db 127 LLNLWCKGRGLICYTSVKWNET 148  
 RESULT 15  
 AAY77369  
 ID AAY77369 standard; protein; 245 AA.  
 XX AC AAY77369;  
 XX XX  
 XX DT 22-MAY-2000 (first entry)  
 XX DE HIV-1 group O pGO-9PL-encoded truncated env gp41 protein, SEQ ID NO:48.  
 XX HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;  
 KW immunoassay; positive control; affinity purification; therapeutic;  
 KW Escherichia coli; antigen; synthetic gene construction; mutagen;  
 KW deletion mutation.  
 XX Human immunodeficiency virus 1; group O isolate HAM112.  
 OS

```

OS Synthetic.
XX WO200004383-A2.
XX PD 27-JAN-2000.
XX PF 09-JUL-1999; 99WO-US015469.
XX PR 14-JUL-1998; 98US-00115171.
XX PA (ABBO ) ABBOTT LAB.
XX PI Scheffel JW, Hackett JR, Tyner JD, Hickman RK;
XX DR WPI: 2000-171290/15.
XX DR N-PSDB; AA290280.
XX
PT Novel monoclonal antibodies useful as positive control reagent for
PT detecting human immunodeficiency virus infections and diagnosing,
PT evaluating or prognosing viral disease.
XX
PS Example 3; Page 120-121; 148pp; English.
XX
CC The invention relates to anti-HIV-1 group O monoclonal antibodies, which
CC may be used as positive control reagents in immunoassays to detect and
CC differentiate HIV-1 infections. The invention also encompasses a
CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,
CC which has no more than 15% cross reactivity to a corresponding antigen
CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of
CC using a monoclonal antibody as a positive control reagent in an
CC immunoassay for the detection of anti HIV-1 group O antibodies. The
CC monoclonal antibodies are useful as positive control reagents in
CC immunoassays capable of detecting anti-HIV-1 group O antibodies. Such
CC immunoassays involve coupling a monoclonal antibody with HIV group-1
CC antigen and detecting the antigen-antibody complex. The monoclonal
CC antibodies of the invention would be used to ensure that the reagents
CC provided to detect HIV-1 group O antibody were performing properly. The
CC monoclonal antibodies may also can be immobilised on a matrix and used
CC for affinity purification of specific HIV-1 group O-derived proteins from
CC cell cultures or biological tissues. The monoclonal antibodies can also
CC be used for generating chimeric antibodies for therapeutic use. Different
CC synthetic, recombinant or purified antibodies which identify different
CC epitopes of HIV antigens can be used in combination in assay to diagnose,
CC evaluate, or prognosticate HIV disease condition. The monoclonal
CC antibodies are also useful for differentiating HIV-1 Group O antigens
CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent
CC recombinant HIV-1 group O env antigens encoded by the synthetic genes
CC AA290280-Z90286. The recombinant HIV-1 env proteins contain various
CC deletions relative to the native HAM112 isolate env protein (AAY77376).
CC The recombinant HIV-1 group O antigens were purified and used to screen
CC hybridoma cultures
XX
SQ Sequence 245 AA;
Query Match 91.3%; Score 116; DB 3; Length 245;
Best Local Similarity 81.8%; Pred. No. 8.5e-08;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRRLVCYTSVQNET 22
Db 127 LLSLWGCRRLVCYTSVQNET 148
||:||||:||||:||||
||:||||:||||:||||

Search completed: May 7, 2004, 17:42:42
Job time : 42.8739 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 12.1092 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-2

Perfect score: 127

Sequence: 1 LLSLMGCRGLVCYTSVQNNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*  
1: /cgn2.6/prodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2.6/prodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2.6/prodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2.6/prodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2.6/prodata/2/iaa/PCUS COMB.pcp.\*  
6: /cgn2.6/prodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	91.3	113	4	US-09-462-917A-40
2	116	91.3	215	2	US-08-912-129A-58
3	116	91.3	245	2	US-08-912-129A-48
4	116	91.3	373	2	US-08-912-129A-52
5	116	91.3	460	2	US-08-912-129A-60
6	116	91.3	490	2	US-08-912-129A-50
7	116	91.3	618	2	US-08-912-129A-54
8	116	91.3	715	4	US-09-462-917A-134
9	116	91.3	873	2	US-08-912-129A-61
10	115	90.6	37	4	US-08-817-441-94
11	115	90.6	104	4	US-08-817-441-100
12	113	89.0	113	4	US-09-462-917A-28
13	113	89.0	116	4	US-09-462-917A-20
14	113	89.0	356	1	US-08-602-713-12
15	113	89.0	356	3	US-08-989-493-12
16	113	89.0	356	4	US-09-610-271-12
17	112	88.2	37	4	US-08-817-441-86
18	112	88.2	113	4	US-09-462-917A-2
19	112	88.2	113	4	US-09-462-917A-18
20	112	88.2	117	4	US-09-462-917A-6
21	112	88.2	200	3	US-08-965-056-104
22	112	88.2	862	4	US-09-206-551-15
23	111	87.4	23	4	US-08-817-441-30
24	111	87.4	23	4	US-09-462-917A-137
25	111	87.4	40	3	US-08-894-699-39
26	111	87.4	40	3	US-09-444-410-39
27	111	87.4	41	3	US-08-894-699-67

28	111	87.4	41	3	US-09-444-410-67	Sequence 67, Appl
29	111	87.4	115	4	US-09-462-917A-24	Sequence 24, Appl
30	111	87.4	116	4	US-09-462-917A-22	Sequence 22, Appl
31	110	86.6	23	4	US-09-462-917A-95	Sequence 95, Appl
32	110	86.6	33	3	US-09-433-428D-8	Sequence 8, Appl
33	110	86.6	40	3	US-08-894-699-41	Sequence 41, Appl
34	110	86.6	40	3	US-08-894-699-42	Sequence 42, Appl
35	110	86.6	40	3	US-09-444-410-41	Sequence 41, Appl
36	110	86.6	40	3	US-09-444-410-42	Sequence 42, Appl
37	110	86.6	110	4	US-09-462-917A-30	Sequence 30, Appl
38	110	86.6	113	4	US-09-462-917A-10	Sequence 10, Appl
39	110	86.6	113	4	US-09-462-917A-12	Sequence 12, Appl
40	109	85.8	40	3	US-08-894-699-37	Sequence 37, Appl
41	109	85.8	40	3	US-09-444-410-37	Sequence 37, Appl
42	108	85.0	24	4	US-09-462-917A-92	Sequence 92, Appl
43	108	85.0	40	3	US-08-894-699-40	Sequence 40, Appl
44	108	85.0	40	3	US-09-444-410-40	Sequence 40, Appl
45	107	84.3	23	3	US-09-433-428D-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1

US-09-462-917A-40  
; Sequence 40, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS-014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; NAME/KEY: Misc feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-09-462-917A-40

Query Match 91.3%; Score 116; DB 4; Length 113;  
Best Local Similarity 81.8%; Pred. No. 6.1e-10;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQNNET 22

Db 39 LLSLMGCRGLVCYTSVQNNKT 60

RESULT 2

US-08-912-129A-58  
; Sequence 58, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.  
; APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.  
; APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.  
; APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.  
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION

NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckerts, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 215 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-912-129A-58

Query Match 91.3%; Score 116; DB 2; Length 215;  
Best Local Similarity 81.8%; Pred. No. 1.2e-09;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||:  
DB 127 LLLWGCKGRLCYTSVKWNET 148

## RESULT 3

US-08-912-129A-48  
Sequence 48, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
APPLICANT: DEVARE, SUSHIL G.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckerts, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 245 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-912-129A-48

Query Match 91.3%; Score 116; DB 2; Length 245;  
Best Local Similarity 81.8%; Pred. No. 1.4e-09;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||:  
DB 127 LLLWGCKGRLCYTSVKWNET 148

## RESULT 4

US-08-912-129A-52  
Sequence 52, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
APPLICANT: DEVARE, SUSHIL G.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Danckerts, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623

TELEX:  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-912-129A-52

Query Match 91.3%; Score 116; DB 2; Length 373;  
Best Local Similarity 81.8%; Pred. No. 2.1e-09;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
DB 127 LLNLWGCKGRLLCYTSVKWNET 148

RESULT 5  
US-08-912-129A-60  
; Sequence 60, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.  
; APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.  
; APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.  
; APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.  
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS (Windows 95)  
; SOFTWARE: Microsoft Word (ASCII format output)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,129A  
; FILING DATE: 15-AUG-1997  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dancckers, Andreas M.  
; REGISTRATION NUMBER: 32,652  
; REFERENCE/DOCKET NUMBER: 6109.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847-937-9803  
; TELEFAX: 847-938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 460 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
US-08-912-129A-60

Query Match 91.3%; Score 116; DB 2; Length 460;  
Best Local Similarity 81.8%; Pred. No. 2.7e-09;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
DB 372 LLNLWGCKGRLLCYTSVKWNET 393

RESULT 6  
US-08-912-129A-50  
; Sequence 50, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.  
; APPLICANT: HICKMAN, ROBERT K.  
; APPLICANT: VARITEK, VINCENT A. JR.  
; APPLICANT: NECKLAWS, ELIZABETH A.  
; APPLICANT: GOLDEN, ALAN M.  
; APPLICANT: BRENNAN, CATHERINE A.  
; APPLICANT: DEVARE, SUSHIL G.  
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS (Windows 95)  
; SOFTWARE: Microsoft Word (ASCII format output)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,129A  
; FILING DATE: 15-AUG-1997  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dancckers, Andreas M.  
; REGISTRATION NUMBER: 32,652  
; REFERENCE/DOCKET NUMBER: 6109.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847-937-9803  
; TELEFAX: 847-938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-912-129A-50

Query Match 91.3%; Score 116; DB 2; Length 490;  
Best Local Similarity 81.8%; Pred. No. 2.9e-09;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
DB 372 LLNLWGCKGRLLCYTSVKWNET 393

RESULT 7  
US-08-912-129A-54  
; Sequence 54, Application US/08912129A  
; Patent No. 5922533  
; GENERAL INFORMATION:  
; APPLICANT: VALLARI, ANADRUZELA S.  
; APPLICANT: HACKETT, JOHN JR.



US-08-817-441-94  
; Sequence 94, Application US/08817441  
; Patent No. 6399294  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; APPLICANT: CLAVEL, FRANCOISE  
; APPLICANT: BORMAN, ANDREW  
; APPLICANT: QUILLIENT, CAROLINE  
; APPLICANT: GUETARD, DENISE  
; APPLICANT: MONTAGNIER, LUC  
; APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; TITLE OF INVENTION: SUBTYPE) ANTIGENS  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,441  
; FILING DATE: 11-JUL-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03260.6005-00000  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 94:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-817-441-94  
Query Match 90.6%; Score 115; DB 4; Length 37;  
Best Local Similarity 81.8%; Pred. No. 2.6e-10;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
Db 14 LLLNLWGCGRKAICYTSVQWNET 35  
RESULT 11  
US-08-817-441-100  
; Sequence 100, Application US/08817441  
; Patent No. 6399294  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; APPLICANT: CLAVEL, FRANCOISE  
; APPLICANT: BORMAN, ANDREW  
; APPLICANT: QUILLIENT, CAROLINE  
; APPLICANT: GUETARD, DENISE  
; APPLICANT: MONTAGNIER, LUC  
; APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; TITLE OF INVENTION: SUBTYPE) ANTIGENS  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,441  
; FILING DATE: 11-JUL-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03260.6005-00000  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 94:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-817-441-100  
; APPLICANT: QUILLIENT, CAROLINE  
; APPLICANT: GUETARD, DENISE  
; APPLICANT: MONTAGNIER, LUC  
; APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE  
; APPLICANT: COHEN, JACQUES  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
; TITLE OF INVENTION: SUBTYPE) ANTIGENS  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,441  
; FILING DATE: 11-JUL-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03260.6005-00000  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 100:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 104 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-817-441-100  
Query Match 90.6%; Score 115; DB 4; Length 104;  
Best Local Similarity 81.8%; Pred. No. 7.8e-10;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LLSLWGCGRGLVCYTSVQWNET 22  
Db 45 LLLNLWGCGRKAICYTSVQWNET 66  
RESULT 12  
US-09-462-917A-28  
; Sequence 28, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS.014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20

```
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
US-09-462-917A-28

Query Match      89.0%; Score 113; DB 4; Length 113;
Best Local Similarity 81.8%; Pred. No. 1.7e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQNNET 22
   ||:||||:||||:||||:||||:
Db 39 LLNLWGCKRGLVCYTSVQNNMT 60

RESULT 13
US-09-462-917A-20
; Sequence 20, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NFUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
US-09-462-917A-20

Query Match      89.0%; Score 113; DB 4; Length 116;
Best Local Similarity 81.8%; Pred. No. 1.7e-09;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQNNET 22
   ||:||||:||||:||||:||||:
Db 39 LLNLWGCKRGLVCYTSVRWNKT 60

RESULT 14
US-08-602-713-12
; Sequence 12, Application US/08602713
; Patent No. 5798205
; GENERAL INFORMATION:
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
; APPLICANT: G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;
; APPLICANT: Zekeng, L opold Achengu
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,713
; FILING DATE: 16-FEBRUARY-1996
; APPLICATION NUMBER: DE 195 05 262
; FILING DATE: 16-FEBRUARY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6162631man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LEDER 203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
```

```
; APPLICATION NUMBER: US/08/602,713
; FILING DATE: 16-FEBRUARY-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 05 262
; FILING DATE: 16-FEBRUARY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5798205man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LEDER 203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: Protein
; FRAGMENT TYPE: internal
US-08-602-713-12

Query Match      89.0%; Score 113; DB 1; Length 356;
Best Local Similarity 77.3%; Pred. No. 5.7e-09;
Matches 17; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQNNET 22
   ||:||||:||||:||||:||||:
Db 322 LLNLWGCKRGLVCYTSVRWNKT 343

RESULT 15
US-08-989-493-12
; Sequence 12, Application US/08989493
; Patent No. 6162631
; GENERAL INFORMATION:
; APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
; APPLICANT: G rtler, Lutz G.; Eberle, Josef; Kaptue, Lazare;
; APPLICANT: Zekeng, L opold Achengu
; TITLE OF INVENTION: Retrovirus From The HIV Group And Its Use
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,713
; FILING DATE: 16-FEBRUARY-1996
; APPLICATION NUMBER: DE 195 05 262
; FILING DATE: 16-FEBRUARY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6162631man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LEDER 203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
```

; LENGTH: 356 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: Protein  
; FRAGMENT TYPE: internal  
; US-08-989-493-12

Query Match 89.0%; Score 113; DB 3; Length 356;  
Best Local Similarity 77.3%; Pred. NO. 5.7e-09;  
Matches 17; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSLWGCRGLVCYTSVQWNET 22  
||:||||:|:||||:||||  
Db 322 LLNLWGCRGLVCYTSVQWNET 343

Search completed: May 7, 2004, 17:53:25  
Job time : 13.1092 secs

SECRET



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 30.7815 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-2  
Perfect score: 127  
Sequence: 1 LLSLMGCRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pbp:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pbp:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pbp:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pbp:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pbp:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTU5\_PUBCOMB.pbp:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pbp:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pbp:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pbp:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pbp:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pbp:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pbp:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pbp:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pbp:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pbp:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pbp:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pbp:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pbp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	91.3	113	14	US-10-320-786-40
2	116	91.3	215	8	US-08-911-824-58
3	116	91.3	245	8	US-08-911-824-48
4	116	91.3	281	8	US-08-911-824-120
5	116	91.3	373	8	US-08-911-824-52
6	116	91.3	460	8	US-08-911-824-60
7	116	91.3	488	8	US-08-911-824-95
8	116	91.3	490	8	US-08-911-824-50
9	116	91.3	526	8	US-08-911-824-97
10	116	91.3	618	8	US-08-911-824-54
11	116	91.3	706	8	US-08-911-824-93
12	116	91.3	715	14	US-10-320-786-134
13	116	91.3	736	8	US-08-911-824-91
14	116	91.3	873	8	US-08-911-824-61
15	115	90.6	37	14	US-10-026-741-94

16	115	90.6	104	14	US-10-026-741-100	Sequence 100, Appl
17	113	89.0	113	14	US-10-320-786-28	Sequence 28, Appl
18	113	89.0	116	14	US-10-320-786-20	Sequence 20, Appl
19	113	89.0	356	14	US-10-357-400-12	Sequence 12, Appl
20	112	88.2	37	14	US-10-026-741-86	Sequence 86, Appl
21	112	88.2	113	14	US-10-320-786-2	Sequence 2, Appl
22	112	88.2	113	14	US-10-320-786-18	Sequence 18, Appl
23	112	88.2	117	14	US-10-320-786-6	Sequence 6, Appl
24	112	88.2	200	9	US-09-854-816-104	Sequence 104, Appl
25	112	88.2	862	14	US-10-369-294-15	Sequence 15, Appl
26	111	87.4	23	14	US-10-026-741-30	Sequence 30, Appl
27	111	87.4	23	14	US-10-320-786-137	Sequence 137, Appl
28	111	87.4	115	14	US-10-320-786-24	Sequence 24, Appl
29	111	87.4	116	14	US-10-320-786-22	Sequence 22, Appl
30	110	86.6	23	14	US-10-320-786-95	Sequence 95, Appl
31	110	86.6	110	14	US-10-320-786-30	Sequence 30, Appl
32	110	86.6	113	14	US-10-320-786-10	Sequence 10, Appl
33	110	86.6	113	14	US-10-320-786-12	Sequence 12, Appl
34	108	85.0	24	14	US-10-320-786-92	Sequence 92, Appl
35	107	84.3	24	14	US-10-320-786-138	Sequence 138, Appl
36	106	83.5	23	14	US-10-320-786-102	Sequence 102, Appl
37	106	83.5	35	14	US-10-026-741-101	Sequence 101, Appl
38	106	83.5	37	14	US-10-026-741-7	Sequence 7, Appl
39	106	83.5	37	14	US-10-026-741-90	Sequence 90, Appl
40	106	83.5	110	14	US-10-320-786-38	Sequence 38, Appl
41	106	83.5	351	14	US-10-026-741-47	Sequence 47, Appl
42	106	83.5	877	14	US-10-026-741-102	Sequence 102, Appl
43	105	82.7	23	14	US-10-320-786-91	Sequence 91, Appl
44	105	82.7	23	14	US-10-320-786-99	Sequence 99, Appl
45	105	82.7	23	14	US-10-320-786-100	Sequence 100, Appl

ALIGNMENTS

RESULT 1  
US-10-320-786-40  
; Sequence 40, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362, 0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-10-320-786-40

Query Match 91.3%; Score 116; DB 14; Length 113;  
Best Local Similarity 81.8%; Pred. No. 2.1e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLSLMGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||:  
Db 39 LLSLMGCRGLVCYTSVQWNET 60

## RESULT 2

US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 91.3%; Score 116; DB 8; Length 215;  
Best Local Similarity 81.8%; Pred. No. 3.8e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
Db 127 LLNLWGCKGRLLCYTSVRKWNET 148

## RESULT 3

US-08-911-824-48  
; Sequence 48, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 48  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL  
US-08-911-824-48

Query Match 91.3%; Score 116; DB 8; Length 245;  
Best Local Similarity 81.8%; Pred. No. 4.2e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
Db 127 LLNLWGCKGRLLCYTSVRKWNET 148

## RESULT 4

US-08-911-824-60

US-08-911-824-120  
; Sequence 120, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 120  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL  
US-08-911-824-120

Query Match 91.3%; Score 116; DB 8; Length 281;  
Best Local Similarity 81.8%; Pred. No. 4.8e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
Db 127 LLNLWGCKGRLLCYTSVRKWNET 148

## RESULT 5

US-08-911-824-52  
; Sequence 52, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 52  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL  
US-08-911-824-52

Query Match 91.3%; Score 116; DB 8; Length 373;  
Best Local Similarity 81.8%; Pred. No. 6.3e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22  
||:||||:||||:||||:||||  
Db 127 LLNLWGCKGRLLCYTSVRKWNET 148

## RESULT 6

US-08-911-824-60

; Sequence 60, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 60

; LENGTH: 460

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS

US-08-911-824-60

Query Match 91.3%; Score 116; DB 8; Length 460;  
Best Local Similarity 81.8%; Pred. No. 7.7e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22

||:||||:||||:||||:||||:||||

Db 372 LLNLWGCKRLICYTSVKWNET 393

#### RESULT 7

US-08-911-824-95

; Sequence 95, Application US/08911824  
; Publication No. US20030004323A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1

; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 95

; LENGTH: 488

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-14PL

US-08-911-824-95

Query Match 91.3%; Score 116; DB 8; Length 488;  
Best Local Similarity 81.8%; Pred. No. 8.1e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22

||:||||:||||:||||:||||:||||

Db 127 LLNLWGCKRLICYTSVKWNET 148

#### RESULT 8

US-08-911-824-50

; Sequence 50, Application US/08911824

; Publication No. US20030004323A1  
; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1

; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 50

; LENGTH: 490

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS

US-08-911-824-50

Query Match 91.3%; Score 116; DB 8; Length 490;  
Best Local Similarity 81.8%; Pred. No. 8.1e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22

||:||||:||||:||||:||||:||||

Db 372 LLNLWGCKRLICYTSVKWNET 393

#### RESULT 9

US-08-911-824-97

; Sequence 97, Application US/08911824  
; Publication No. US20030004323A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1

; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 97

; LENGTH: 526

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS

US-08-911-824-97

Query Match 91.3%; Score 116; DB 8; Length 526;  
Best Local Similarity 81.8%; Pred. No. 8.7e-08;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22

||:||||:||||:||||:||||:||||

Db 372 LLNLWGCKRLICYTSVKWNET 393

#### RESULT 10

US-08-911-824-54

; Sequence 54, Application US/08911824  
; Publication No. US20030004323A1

```
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11CKS
US-08-911-824-54

Query Match          91.3%; Score 116; DB 8; Length 618;
Best Local Similarity 81.8%; Pred. No. 1e-07;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22
||:||||:||||:||||:||||
Db 372 LNLWGCKGRLLCYTSVKWNET 393

RESULT 11
US-08-911-824-93
; Sequence 93, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-13CKS
US-08-911-824-93

Query Match          91.3%; Score 116; DB 8; Length 706;
Best Local Similarity 81.8%; Pred. No. 1.1e-07;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22
||:||||:||||:||||:||||
Db 618 LNLWGCKGRLLCYTSVKWNET 639

RESULT 12
US-10-320-786-134
; Sequence 134, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marleen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Human
; OTHER INFORMATION: Encodes recombinant protein pGO-12CKS
US-10-320-786-134

Query Match          91.3%; Score 116; DB 14; Length 715;
Best Local Similarity 86.4%; Pred. No. 1.2e-07;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22
||:||||:||||:||||:||||
Db 628 LNLWGCKGRLLVCYTSVQWNET 649

RESULT 13
US-08-911-824-91
; Sequence 91, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-12CKS
US-08-911-824-91

Query Match          91.3%; Score 116; DB 8; Length 736;
Best Local Similarity 81.8%; Pred. No. 1.2e-07;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22
||:||||:||||:||||:||||
Db 618 LNLWGCKGRLLCYTSVKWNET 639

RESULT 14
US-08-911-824-61
; Sequence 61, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
```



Conrad ... ..

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 9.2437 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-2

Perfect score: 127  
Sequence: 1 LLSLWGCRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	115	90.6	104	2 S52930	GP41 ENV protein -
2	112	88.2	863	2 A53034	gag polyprotein -
3	106	83.5	877	2 S49197	envelope protein p
4	89	70.1	854	1 VCLJJI	env polyprotein pr
5	86	67.7	357	2 S21990	envelope protein g
6	86	67.7	358	2 S22002	envelope protein g
7	85	66.9	358	2 S22000	envelope protein g
8	85	66.9	358	2 S70417	envelope protein g
9	84	66.1	443	2 C41621	env polyprotein P
10	84	66.1	853	2 S54384	envelope polyprote
11	84	66.1	855	1 VCLJZR	env polyprotein pr
12	83	65.4	357	2 S22006	envelope protein g
13	83	65.4	357	2 S21994	envelope protein g
14	83	65.4	357	2 S22004	envelope protein g
15	83	65.4	357	2 S21996	envelope protein g
16	83	65.4	357	2 S21992	envelope protein g
17	83	65.4	358	2 S21998	envelope protein g
18	83	65.4	445	2 A41621	env polyprotein M
19	83	65.4	454	2 B41621	env polyprotein D
20	83	65.4	843	1 H44001	env polyprotein pr
21	83	65.4	847	2 T09448	envelope glycoprot
22	83	65.4	847	2 S13289	env protein - huma
23	83	65.4	852	1 VCLJBR	env polyprotein -
24	83	65.4	852	1 T12016	envelope glycoprot
25	83	65.4	854	2 S13288	env protein - huma
26	83	65.4	855	1 VCLJAJ2	env polyprotein pr
27	83	65.4	856	1 VCLJH3	env polyprotein pr
28	83	65.4	856	1 VCLJVL	env polyprotein pr
29	83	65.4	856	1 VCLJ3W	env polyprotein pr

30	83	65.4	861	1 VCLJLV	env polyprotein pr
31	83	65.4	861	1 VCLJSC	env polyprotein pr
32	81	63.8	856	1 A44963	env polyprotein pr
33	81	63.8	859	1 VCLJMN	env polyprotein pr
34	81	63.8	868	1 VCLJH4	env polyprotein -
35	80	63.0	846	1 VCLJND	env polyprotein pr
36	77	60.6	729	1 VCLJXX	env polyprotein pr
37	77	60.6	861	1 VCLJKB	env polyprotein pr
38	76	59.8	786	2 S28084	env polyprotein -
39	76	59.8	851	2 S33985	env polyprotein -
40	74	58.3	859	2 T01672	envelope polyprote
41	72.5	57.1	855	2 A45713	Env transmembrane
42	68.5	53.9	859	1 VCLJST	env polyprotein pr
43	68.5	53.9	859	2 S24571	env protein - huma
44	68.5	53.9	885	2 S04322	env polyprotein -
45	68.5	53.9	886	2 T11555	env protein - simi

ALIGNMENTS

RESULT 1

S52930

GP41 ENV protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C:Accession: S52930

R:Cohen, J.H.M.; Guetard, D.; Philbert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; d

Submitted to the EMBL Data Library, January 1995

A:Description: A novel HIV1-O strain illustrates the diversity of the O group.

A:Reference number: S52929

A:Accession: S52930

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-104 <COH>

A:Cross-references: EMBL:X84328; NID:G695526; PIDN:CAAS9066.1; PID:G695527

C:Superfamily: type E retrovirus env polyprotein

Query Match 90.6%; Score 115; DB 2; Length 104;

Best Local Similarity 81.8%; Pred. No. 1.3e-09;

Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22

Db 45 LLSLWGCRGLVCYTSVQWNET 66

|||||: |||||

|||||: |||||

RESULT 2

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborgh, B.; van der Groen,

J. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African h

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-863 <VANS>

A:Cross-references: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 88.2%; Score 112; DB 2; Length 863;

Best Local Similarity 81.8%; Pred. No. 2.2e-08;

Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCRGLVCYTSVQWNET 22

Db 594 LLSLWGCRGLVCYTSVQWNET 615

|||||: |||||

|||||: |||||





```
F;424-443/Domain: transmembrane #status predicted <TM>
F;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: C
Query Match          66.1%; Score 84; DB 2; Length 443;
Best Local Similarity 54.5%; Pred. No. 0.00014;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSTVQWNEN 22
   |||||:::||:||||:
Db 332 LLGIWGCSGKLICTTTPWNSS 353

RESULT 10
S54384
envelope polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C;Accession: S54384
R;Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A;Reference number: S54377
A;Accession: S54384
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-853 <PHE>
A;Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385
C;Superfamily: type E retrovirus env glycoprotein
C;Keywords: polyprotein

Query Match          66.1%; Score 84; DB 2; Length 853;
Best Local Similarity 54.5%; Pred. No. 0.00025;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSTVQWNEN 22
   |||||:::||:||||:
Db 589 LLGIWGCSGKLICTTTPWNSS 610

RESULT 11
VCLJZR
env polyprotein precursor - human immunodeficiency virus Zr-6
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus Zr-6
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: D26192
R;Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Ch
Gene 52, 71-82, 1987
A;Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot
A;Reference number: A26192; UID:87248097; PMID:3036660
A;Accession: D26192
A;Molecule type: DNA
A;Residues: 1-855 <SR>
A;Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45380.1; PID:g329403
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-855/Product: env polyprotein #status predicted <MAT>
F;501-855/Product: exterior membrane glycoprotein #status predicted <EXT>
F;87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,40
Query Match          66.1%; Score 84; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.00025;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSTVQWNEN 22
   |||||:::||:||||:
Db 591 LLGIWGCSGKLICTTTPWNSS 612

RESULT 12
```

```
S22006
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70420; S22006
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70420
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191
A:Experimental source: patient L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00016;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSLMGCGRLVCYTSVQWN 20
   ||:|||||:|:|:|:|:|
Db 93 LLGIWCGSGKLICTTTPWN 112

RESULT 13
S21994
envelope protein gp120/gp41 - human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 27B
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21994; S70421
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
A:Accession: S21994
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140,'X',142-312,'X',314-357 <ST2>
A:Cross-references: EMBL:X61355; NID:g60179
C:Superfamily: type E retrovirus env polyprotein

Query Match          65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00016;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSLMGCGRLVCYTSVQWN 20
   ||:|||||:|:|:|:|:|
Db 93 LLGIWCGSGKLICTTTPWN 112

RESULT 14
S22004
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 4B
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22004; S70419
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
```

```
A:Accession: S22004
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70419
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292,'X',294-357 <ST2>
A:Cross-references: EMBL:X61353; NID:g60188
C:Superfamily: type E retrovirus env polyprotein

Query Match          65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00016;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSLMGCGRLVCYTSVQWN 20
   ||:|||||:|:|:|:|:|
Db 93 LLGIWCGSGKLICTTTPWN 112

RESULT 15
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70422; S21996
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43624.1; PID:g1067129
A:Experimental source: patient 27L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00016;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSLMGCGRLVCYTSVQWN 20
   ||:|||||:|:~|:|:|
Db 93 LLGIWCGSGKLICTTTPWN 112

Search completed: May 7, 2004, 17:51:01
Job time : 10.2437 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 5.54622 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-2  
Perfect score: 127  
Sequence: 1 LLSLMGCRGLVCYTSVQWNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	70.1	854	1 ENV_SIVCZ	P17281 chimpanzee
2	84	66.1	853	1 ENV_HV1Z2	P12487 human immun
3	84	66.1	855	1 ENV_HV1Z6	P04580 human immun
4	83	65.4	843	1 ENV_HV1Y2	P35961 human immun
5	83	65.4	847	1 ENV_HV1S1	P19550 human immun
6	83	65.4	847	1 ENV_HV1W2	P05880 human immun
7	83	65.4	851	1 ENV_HV1B8	P04582 human immun
8	83	65.4	852	1 ENV_HV1B9	P12488 human immun
9	83	65.4	852	1 ENV_HV1S3	P19549 human immun
10	83	65.4	853	1 ENV_HV1MF	P19551 human immun
11	83	65.4	855	1 ENV_HV1A2	P03378 human immun
12	83	65.4	855	1 ENV_HV1OY	P20888 human immun
13	83	65.4	856	1 ENV_HV1B1	P03375 human immun
14	83	65.4	856	1 ENV_HV1H2	P04578 human immun
15	83	65.4	856	1 ENV_HV1LW	P04578 human immun
16	83	65.4	856	1 ENV_HV1PV	P03376 human immun
17	83	65.4	856	1 ENV_HV1SC	P05878 human immun
18	83	65.4	856	1 ENV_HV1W1	P31872 human immun
19	83	65.4	861	1 ENV_HV1BR	P03377 human immun
20	83	65.4	865	1 ENV_HV1RH	P04579 human immun
21	83	65.4	867	1 ENV_HV1J3	P12489 human immun
22	81	63.8	848	1 ENV_HV1JR	P20871 human immun
23	81	63.8	856	1 ENV_HV1H3	P04624 human immun
24	81	63.8	856	1 ENV_HV1MN	P05877 human immun
25	81	63.8	856	1 ENV_HV1ZH	P05881 human immun
26	81	63.8	868	1 ENV_HV1C4	P05879 human immun
27	80	63.0	846	1 ENV_HV1ND	P18799 human immun
28	77	60.6	853	1 ENV_HV1EL	P04581 human immun
29	77	60.6	861	1 ENV_HV1KB	P31819 human immun
30	77	60.6	863	1 ENV_HV1Z8	P05882 human immun
31	76	59.8	821	1 ENV_SIVGB	P22380 simian immu
32	74	58.3	859	1 ENV_HV1MA	P04583 human immun
33	68.5	53.9	856	1 ENV_HV2N2	P05883 human immun

RESULT 1

ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheyner R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1.;"			
RL	Nature 345:356-359(1990).			
CC	-!- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X52154; CAA36407.1; -.			
DR	FIR; S09990; VCLJSL.			
DR	HIV; X52154; ENVSCPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	31	500	TRANSMEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	POTENTIAL.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. .) (POTENTIAL).

ALIGNMENTS

34	68.5	53.9	859	1 ENV_HV2D2	P15831 human immun
35	68.5	53.9	885	1 ENV_SIVS4	P12492 simian immu
36	67	52.8	857	1 ENV_HV2KR	Q74126 human immun
37	65.5	51.6	712	1 ENV_HV2S2	P32536 human immun
38	65.5	51.6	846	1 ENV_HV2SB	P12449 human immun
39	65.5	51.6	859	1 ENV_HV2ST	P20872 human immun
40	65	51.2	882	1 ENV_SIVM1	P05885 simian immu
41	64.5	50.8	859	1 ENV_HV2CA	P24105 human immun
42	64	50.4	380	1 ENV_SIVM2	P08810 simian immu
43	64	50.4	865	1 ENV_SIVAT	P05886 simian immu
44	64	50.4	881	1 ENV_SIVMK	P05884 simian immu
45	63	49.6	851	1 ENV_HV2D1	P17755 human immun

```

FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 70.1%; Score 89; DB 1; Length 854;
Best Local Similarity 59.1%; Pred. No. 4.4e-06;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLSLWGCGRGLVCYTSVQVNET 22
Db 582 ILGLWGCGRGLVCYTSVQVNET 603

RESULT 2
ENV_HV122
ID ENV_HV122 STANDARD; PRT; 853 AA.
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV5226.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```

FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCAFF7008 CRC64;

Query Match 66.1%; Score 84; DB 1; Length 853;
Best Local Similarity 54.5%; Pred. No. 2.6e-05;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLSLWGCGRGLVCYTSVQVNET 22
Db 589 LLGLWGCGRGLVCYTSVQVNET 610

RESULT 3
ENV_HV126
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.

```





FT DISULFID 372 435 BY SIMILARITY.  
FT FT CARBOHYD 379 408 BY SIMILARITY.  
FT FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 847 AA; 96466 MW; CD1E33D73A5BCAE CRC64;  
  
Query Match 65.4%; Score 83; DB 1; Length 847;  
Best Local Similarity 60.0%; Pred. No. 3.7e-05;  
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 LLSLMGCRGLVCYTSVQWN 20  
||:||||:|:|:|  
Db 583 LLGWCSCGKLICTTTVPWN 602  
  
RESULT 7  
ENV HV1B8 STANDARD; PRT; 851 AA.  
AC P04582;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (BH8 isolate). (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11694;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumanis K., Ivanoff L., Pettway S.R., Pearson M.L., Lautenberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C., Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; K02011; AAA44661.1; -.  
DR PDB; 1DDH; 13-JAN-99.  
DR PDB; 1HHG; 31-OCT-93.  
DR PDB; 1003; 02-JAN-00.  
DR PDB; 1S2T; 24-DEC-97.  
DR HIV; K02011; ENV\$B8.  
DR GlycosuiteDB; P04582; -.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;  
3D-structure.  
FT SIGNAL 1 30  
FT CHAIN 31 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 507 851 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 205 BY SIMILARITY.  
FT DISULFID 126 196 BY SIMILARITY.  
FT DISULFID 131 157 BY SIMILARITY.  
FT DISULFID 218 247 BY SIMILARITY.  
FT DISULFID 228 239 BY SIMILARITY.  
FT DISULFID 296 331 BY SIMILARITY.  
FT DISULFID 378 440 BY SIMILARITY.  
FT DISULFID 385 413 BY SIMILARITY.  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;  
  
Query Match 65.4%; Score 83; DB 1; Length 851;  
Best Local Similarity 60.0%; Pred. No. 3.7e-05;  
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 LLSLMGCRGLVCYTSVQWN 20  
||:||||:|:|:|  
Db 587 LLGWCSCGKLICTTTVPWN 606  
  
RESULT 8  
ENV HV1B8 STANDARD; PRT; 852 AA.  
AC P12488;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane









-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; K02007; AAB59882.1; -;  
PIR; A03976; VCLJ2A.  
HIV; K02007; ENVSSSF2.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
FT SIGNAL 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 208 BY SIMILARITY.  
FT DISULFID 125 199 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 221 250 BY SIMILARITY.  
FT DISULFID 221 242 BY SIMILARITY.  
FT DISULFID 229 333 BY SIMILARITY.  
FT DISULFID 380 442 BY SIMILARITY.  
FT DISULFID 387 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 65.4%; Score 83; DB 1; Length 855;  
Best Local Similarity 60.0%; Pred. No. 3.8e-05;  
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWN 20

Db 591 LLGIWGCGRGLVCYTSVQWN 610

RESULT 12

ENV\_HV10Y  
ID ENV\_HV10Y STANDARD; PRT; 855 AA.  
AC P20888;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90148544; PubMed=2559749;  
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;  
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese  
individual presenting an atypical western blot.";  
RL AIDS 3:707-715(1989).  
CC -I- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A  
HEALTHY GABONESE INDIVIDUAL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M26727; AAA83397.1; -;  
DR HIV; M26727; ENVSOVI.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
FT SIGNAL 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 210 BY SIMILARITY.  
FT DISULFID 125 201 BY SIMILARITY.  
FT DISULFID 130 162 BY SIMILARITY.  
FT DISULFID 223 252 BY SIMILARITY.  
FT DISULFID 233 244 BY SIMILARITY.  
FT DISULFID 301 335 BY SIMILARITY.  
FT DISULFID 381 442 BY SIMILARITY.  
FT DISULFID 388 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).





FT	CARBOHYD	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA; 96938 MW; 0C241332CF7E6687 CRC64;		

Query Match 65.4%; Score 83; DB 1; Length 856;  
 Best Local Similarity 60.0%; Pred. No. 3.8e-05;  
 Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSLWGRGLVCYTSVQWN 20  
 Db 592 LIGWCSGKLICTTAVPMN 611

Search completed: May 7, 2004, 17:43:54  
 Job time : 5.54622 secs

General Property and  
Business Insurance Co.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 27.916 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-2

Perfect score: 127

Sequence: 1 LLSLMGCRGLVCYTSVQWNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	93.7	209	15	Q9IE66 human immun
2	116	91.3	116	15	O40458 human immun
3	116	91.3	124	15	Q9IHU7 human immun
4	116	91.3	126	15	Q9IHV1 human immun
5	116	91.3	158	15	Q8J3N6 human immun
6	116	91.3	230	15	Q9IEB2 human immun
7	116	91.3	240	15	Q9IE32 human immun
8	116	91.3	536	15	Q9IEE5 human immun
9	116	91.3	879	15	Q9WIU9 human immun
10	116	91.3	880	15	Q8Q7H1 human immun
11	116	91.3	900	15	Q9QNZ8 human immun
12	115	90.6	104	15	Q76163 human immun
13	115	90.6	216	15	Q9IEC7 human immun
14	115	90.6	535	15	Q9IEF2 human immun
15	114	89.8	125	15	Q9IHU8 human immun
16	114	89.8	230	15	Q9IEA9 human immun

17	114	89.8	238	15	Q9DIK1	Q9dik1 human immun
18	114	89.8	529	15	Q9IEE2	Q9iee2 human immun
19	113	89.0	114	15	O40448	O40448 human immun
20	113	89.0	116	15	O40459	O40459 human immun
21	113	89.0	137	15	Q9IHV5	Q9ihv5 human immun
22	113	89.0	155	15	Q8J3Q4	Q8j3q4 human immun
23	113	89.0	172	15	Q9IEB3	Q9ieb3 human immun
24	113	89.0	177	15	Q9IEB0	Q9ieb0 human immun
25	113	89.0	208	15	Q9IEA3	Q9iea3 human immun
26	113	89.0	227	15	Q9IE99	Q9ie99 human immun
27	113	89.0	234	15	Q9IEC2	Q9iec2 human immun
28	113	89.0	342	15	O11942	O11942 human immun
29	113	89.0	418	15	O36547	O36547 human immun
30	113	89.0	502	15	Q9IEE3	Q9iee3 human immun
31	113	89.0	871	15	O57074	O57074 human immun
32	113	89.0	871	15	Q8Q7I2	Q8q7i2 human immun
33	113	89.0	876	15	Q8Q7H3	Q8q7h3 human immun
34	112	88.2	114	15	O40456	O40456 human immun
35	112	88.2	114	15	O40472	O40472 human immun
36	112	88.2	118	15	O40451	O40451 human immun
37	112	88.2	153	15	Q8J3R1	Q8j3r1 human immun
38	112	88.2	213	15	Q9IEC4	Q9iec4 human immun
39	112	88.2	225	15	Q9IEA0	Q9iea0 human immun
40	112	88.2	234	15	Q9IEA6	Q9iea6 human immun
41	112	88.2	532	15	Q9IEF0	Q9ief0 human immun
42	112	88.2	545	15	Q9IED7	Q9ied7 human immun
43	112	88.2	551	15	Q9IEE1	Q9iee1 human immun
44	112	88.2	863	15	O77377	O77377 human immun
45	112	88.2	872	15	Q900Y5	Q900y5 human immun

ALIGNMENTS

RESULT 1

Q9IE66	PRELIMINARY;	PRT;	209 AA.
ID	Q9IE66		
AC	Q9IE66;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Env polypeptide, gp41 (Fragment).		
GN	ENV.		
OS	Human immunodeficiency virus 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BCF107;		
RA	Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,		
RA	Depienne C., Brun-Vezinet F., Dormont D., Simon P.O.;		
RT	"Phylogenetic analysis and subtyping of 47 HIV-1."		
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ238862; CAB96300.1; -		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0019031; C:viral envelope; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR000328; Env_GP41.		
DR	Pfam; PF00517; GP41; 1.		
KW	Transmembrane.		
FT	NON_TER 1		
FT	NON_TER 209		
SQ	SEQUENCE 209 AA; 24116 MW; 91AC9BE2BFE74B83 CRC64;		

Query Match 93.7%; Score 119; DB 15; Length 209;  
Best Local Similarity 86.4%; Pred. No. 2.2e-10;  
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLMGCRGLVCYTSVQWNET 22

|||||:|:|:|:|:|:|:|:|:|

Db 56 LLSLMGCKRGLICYSVKWNET 77

Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSLWGCRGLVCVTSVQNMNET 22  
 ||:||||:||||:||||:||||  
 Db 33 LLNLWGCKGLICYTSVKWNEN 54

RESULT 4  
 Q9IHV1 PRELIMINARY; PRT; 126 AA.  
 ID Q9IHV1  
 AC Q9IHV1  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN GP41.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 [1]  
 RN  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=97ES203;  
 RX MEDLINE=20386754; PubMed=10933623;  
 RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,  
 RA Pientazek D., Schable C., Lal R.B.;  
 RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
 RT type 1 group O";  
 RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
 DR EMBL; AF229233; AAF71910.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON TER 1  
 FT NON TER 126 126  
 FT NON TER 126 126  
 SQ SEQUENCE 126 AA; 15169 MW; 13FB101BCDF0DD CRC64;

Query Match 91.3%; Score 116; DB 15; Length 126;  
 Best Local Similarity 86.4%; Pred. No. 3.8e-10;  
 Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSLWGCRGLVCVTSVQNMNET 22  
 ||:||||:||||:||||:||||  
 Db 40 LLNLWGCKGLICYTSVKWNES 61

RESULT 5  
 Q8J3N6 PRELIMINARY; PRT; 158 AA.  
 ID Q8J3N6  
 AC Q8J3N6  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE GP41 protein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 [1]  
 RN  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=22092513; PubMed=12097573;  
 RA Boutonnet N., Janssens W., Bouton C., Verschelde J.L., Heyndrickx L.,  
 RA Beirnaert E., van der Groen G., Laesters I.;  
 RT "Comparison of predicted scaffold-compatible sequence variation in the  
 RT triple-hairpin structure of human immunodeficiency virus type 1 gp41  
 RT with patient data";  
 RL J. Virol. 76:7595-7606(2002).  
 DR EMBL; AJ428017; CAD20969.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.



```
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 158
FT NON_TER 158
SQ SEQUENCE 158 AA; 19803 MW; 6B8AD855599B3CC8 CRC64;

Query Match          91.3%; Score 116; DB 15; Length 158;
Best Local Similarity 81.8%; Pred. No. 4.8e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 48 LLNLWGCGRGLVCYTSVQWNET 69

RESULT 6
Q9IEB2 PRELIMINARY; PRT; 230 AA.
AC Q9IEB2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF15;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236408; CAB96256.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1
FT NON_TER 230
FT NON_TER 230
SQ SEQUENCE 230 AA; 26705 MW; C1F744F82F51302D CRC64;

Query Match          91.3%; Score 116; DB 15; Length 230;
Best Local Similarity 81.8%; Pred. No. 6.9e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 55 LLTLWGCGRGLVCYTSVQWNET 76

RESULT 7
Q9IE32 PRELIMINARY; PRT; 240 AA.
AC Q9IE32;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF111;
RA Roques P., Robertson D., Sousquiere S., Damond F., Maucelere P.,
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236408; CAB96256.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1
FT NON_TER 230
FT NON_TER 230
SQ SEQUENCE 230 AA; 26705 MW; C1F744F82F51302D CRC64;

Query Match          91.3%; Score 116; DB 15; Length 230;
Best Local Similarity 81.8%; Pred. No. 6.9e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 55 LLTLWGCGRGLVCYTSVQWNET 76

RESULT 7
Q9IE32 PRELIMINARY; PRT; 240 AA.
AC Q9IE32;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF111;
RA Roques P., Robertson D., Sousquiere S., Damond F., Maucelere P.,
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
```

```
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243365; CAB96335.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane.
FT NON_TER 1
FT NON_TER 240
FT NON_TER 240
SQ SEQUENCE 240 AA; 27723 MW; 1E206E8D491A4197 CRC64;

Query Match          91.3%; Score 116; DB 15; Length 240;
Best Local Similarity 81.8%; Pred. No. 7.2e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 60 LLNLWGCGRGLVCYTSVQWNET 81

RESULT 8
Q9IEE5 PRELIMINARY; PRT; 536 AA.
AC Q9IEE5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF15;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133062; CAB96223.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 536
FT NON_TER 536
SQ SEQUENCE 536 AA; 60419 MW; 0F533ABAF08F5552 CRC64;

Query Match          91.3%; Score 116; DB 15; Length 536;
Best Local Similarity 81.8%; Pred. No. 1.6e-09;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRGLVCYTSVQWNET 22
DB 361 LLTLWGCGRGLVCYTSVQWNET 382

RESULT 9
Q9WIU9 PRELIMINARY; PRT; 879 AA.
AC Q9WIU9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP160 precursor.
GN ENV.
OS Human immunodeficiency virus 1.
```

```
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_
RP SEQUENCE FROM N.A.
RX STRAIN=V1686;
RC MEDLINE=99223950; PubMed=10207543;
RA Janssens W., Heyndrickx L., Van der Auwera G., Nkengasong J.,
RA Beirnaert E., Verbecken K., Coppens S., Willems B., Franssen K.,
RA Peeters M., Ndumbe P., Delaporte E., Van der Groen G.;
RT "Interpatient genetic variability of HIV-1 group O.";
RL AIDS 13:41-48(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=V1686;
RC MEDLINE=96426454; PubMed=8828748;
RA Delaporte E., Janssens W., Peeters M., Buve A., Dibanga G.,
RA Ferrer J.L., Ditsambou V., Georges Courbot M.C., Georges A.,
RA Bourgeois A., Samb B., Henzel D., Heyndrickx L., Franssen K.,
RA Van der Groen G., Larouz B., Mbe J.R.;
RT "Epidemiological and Molecular characteristics of HIV infection in
RT Gabon (1986 - 1994).";
RL AIDS 10:903-910(1996).
DR EMBL; X96526; CAA65373.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; AIDS; Coat protein; Glycoprotein; Polyprotein; Signal; Transmembrane.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 524 GP120.
FT CHAIN 132 156 V1.
FT CHAIN 164 200 V2.
FT CHAIN 302 335 V3.
FT CHAIN 395 427 V4.
FT CHAIN 458 475 V5.
FT CHAIN 525 879 GP41.
SQ SEQUENCE 879 AA; 99099 MW; BC6422ADDFF1A3409 CRC64;

Query Match
Best Local Similarity 91.3%; Score 116; DB 15; Length 879;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRLVCYTSVQWNET 22
||:||||:||||:||||:||||:
Db 605 LLNLWGCKGRLCYTSVQWNET 626

RESULT 10
QY Q8Q7H1 PRELIMINARY; PRT; 880 AA.
AC Q8Q7H1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=98CMA407;
RC MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Nganso C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383250; AAL98872.1; -.
```

```
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 880 AA; 99784 MW; F50D9358BED301EE CRC64;

Query Match
Best Local Similarity 91.3%; Score 116; DB 15; Length 880;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSLWGCGRLVCYTSVQWNET 22
||:||||:||||:||||:||||:
Db 608 LLNLWGCKGRLCYTSVQWNET 629

RESULT 11
QY Q9QN28 PRELIMINARY; PRT; 900 AA.
AC Q9QN28;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ENV protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CA-MP645M/O;
RX MEDLINE=99370180; PubMed=10438826;
RA Peeters M., Liegeois F., Torimino N., Bourgeois A., Mpoudi E.,
RA Vergne L., Saman E., Delaporte E., Saragosti S.;
RT "Characterization of a highly replicative intergroup M/O human
RT immunodeficiency virus type 1 recombinant isolated from a Cameroonian
RT patient.";
RL J. Virol. 73:7368-7375(1999).
DR EMBL; AJ239083; CAB53242.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 900 AA; 100967 MW; FDB5E726A2537966 CRC64;

Query Match
Best Local Similarity 86.4%; Score 116; DB 15; Length 900;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSLWGCGRLVCYTSVQWNET 22
||:||||:||||:||||:||||:
Db 628 LLNLWGCKGRLCYTSVQWNET 649

RESULT 12
QY Q76163 PRELIMINARY; PRT; 104 AA.
AC Q76163;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 ENV (Fragment).
GN HIV-0 ENV.
OS Human immunodeficiency virus 1.
```

DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)	
DE	Env polypeptide (Fragment).	
GN	ENV.	
OS	Human immunodeficiency virus 1.	
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11676;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BCF03;	
RA	Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,	
RA	Deleppe C., Brun-Vezinet F., Dormont D.;	
RT	"HIV-1 group O phylogenetic analysis of C2-gp41 region.";	
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AJ133055; CAB96216.1; -	
DR	GO: GO:0016021; C:integral to membrane; IEA.	
DR	GO: GO:0019028; C:viral capsid; IEA.	
DR	GO: GO:0019031; C:viral envelope; IEA.	
DR	GO: GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro: IPR000328; Env GP41.	
DR	InterPro: IPR000777; GP120.	
DR	Pfam: PF00516; GP120; 1.	
DR	Pfam: PF00517; GP41; 1.	
KW	AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.	
FT	NON_TER 1	
FT	NON_TER 535 535	
FT	SEQUENCE 535 AA; 59682 MW; 7234BE8378DD12C5 CRC64;	
SQ		
Query Match	90.6%; Score 115; DB 15; Length 535;	
Best Local Similarity	86.4%; Pred. No. 2.2e-09;	
Matches	19; Conservative 2; Mismatches 1; Indels 0; Gap	
Qy	1 LLSLWCGRLVCYTSVQWNET 22	
	:     :	
Db	374 LLSLWCGRLVCYTSVKWNET 395	
RESULT 15		
QIUH8		
ID	PRELIMINARY; PRT; 125 AA.	
AC	QIUH8;	
DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)	
DE	Envelope glycoprotein (Fragment).	
GN	GP41.	
OS	Human immunodeficiency virus 1.	
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11676;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=97CM766;	
RX	MEDLINE=20386754; PubMed=10933623;	
RA	Yang C., Gao F., Fonjuncq P.N., Zekeng L., van der Groen G.,	
RA	Pieniazek D., Schable C., Lal R.B.;	
RT	"Phylogenetic analysis of protease and transmembrane regions of HIV	
RT	type 1 group O.";	
RL	AIDS Res. Hum. Retroviruses 16:1075-1081(2000).	
DR	EMBL: AF229236; AAF1913.1; -	
DR	GO: GO:0016021; C:integral to membrane; IEA.	
DR	GO: GO:0019031; C:viral envelope; IEA.	
DR	GO: GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro: IPR000328; Env_GP41.	
DR	Pfam: PF00517; GP41; 1.	
KW	Transmembrane.	
FT	NON_TER 1	
FT	NON_TER 125 125	
FT	SEQUENCE 125 AA; 15108 MW; EA2896C4B25FF342 CRC64;	
SQ		
Query Match	89.8%; Score 114; DB 15; Length 125;	
Best Local Similarity	77.3%; Pred. No. 7.7e-10;	
Matches	17; Conservative 5; Mismatches 0; Indels 0; Gap	

Qy 1 LLSLWGCRGLVCYTISVOWNET 22  
||:||||:||||:||||:||||  
Db 37 LLNLWGCKGRILICYTISVKWNET 58

Search completed: May 7, 2004, 17:49:11  
Job time : 29.916 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 41.8739 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-3

Perfect score: 127  
Sequence: 1 LLSSWGCKRLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	100.0	22	2 AAW80461	Peptide d
2	124	97.6	22	2 AAW80466	Peptide d
3	122	96.1	22	2 AAW80462	Peptide d
4	120	94.5	40	2 AAW07346	Partial s
5	118	92.9	22	2 AAW80463	Peptide d
6	118	92.9	22	2 AAW80460	Peptide d
7	117	92.1	32	2 AAW80469	Peptide d
8	113	89.0	22	2 AAW80465	Peptide d
9	113	89.0	22	2 AAW80464	Peptide d
10	113	89.0	40	2 AAW07352	Partial s
11	113	89.0	113	2 AAY05565	HIV-1 gro
12	113	89.0	200	3 AAY77373	HIV-1 gro
13	113	89.0	215	2 AAY09499	HIV-1 gro
14	113	89.0	215	2 AAY06983	Recombina
15	113	89.0	215	3 AAY77374	HIV-1 gro
16	113	89.0	245	2 AAY09493	HIV-1 gro
17	113	89.0	245	2 AAY06977	Recombina
18	113	89.0	245	3 AAY77369	HIV-1 gro
19	113	89.0	281	2 AAY09507	HIV-1 gro
20	113	89.0	373	2 AAY09495	HIV-1 gro
21	113	89.0	373	2 AAY06979	Recombina
22	113	89.0	460	2 AAY09500	HIV-1 gro
23	113	89.0	460	2 AAY06984	Recombina
24	113	89.0	460	3 AAY77375	HIV-1 gro
25	113	89.0	474	3 AAY77371	HIV-1 gro

26	113	89.0	488	2 AAY09504	HIV-1 Gro
27	113	89.0	490	2 AAY09494	HIV-1 Gro
28	113	89.0	490	2 AAY06978	Recombina
29	113	89.0	490	3 AAY77370	HIV-1 gro
30	113	89.0	526	2 AAY09505	HIV-1 Gro
31	113	89.0	618	2 AAY09496	HIV-1 Gro
32	113	89.0	618	2 AAY06980	Recombina
33	113	89.0	618	3 AAY77372	HIV-1 gro
34	113	89.0	706	2 AAY09503	HIV-1 Gro
35	113	89.0	715	2 AAY05625	HIV-1 Gro
36	113	89.0	736	2 AAY09502	HIV-1 Gro
37	113	89.0	873	2 AAY09501	HIV-1 Gro
38	113	89.0	873	2 AAY06985	Amino aci
39	113	89.0	873	3 AAY77376	Peptide d
40	111	87.4	32	2 AAW80470	Partial s
41	111	87.4	33	3 AAB12212	Peptide d
42	110	86.6	22	2 AAW80459	Peptide d
43	110	86.6	32	2 AAW80471	Peptide d
44	110	86.6	33	3 AAB12231	Partial s
45	110	86.6	40	2 AAW07343	Partial s

ALIGNMENTS

RESULT 1  
AAW80461  
ID AAW80461 standard; peptide; 22 AA.

AC AAW80461;

XX

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)

XX Peptide derived from a conserved sequence of group O human HIV.

XX Group O human immune deficiency virus; HIV; detection; infection.

XX Synthetic.

OS Human immunodeficiency virus 1.

XX WO9845323-A1.

XX 15-OCT-1998.

PD 06-APR-1998; 98WO-FR000691.

PF 09-APR-1997; 97FR-00004356.

PR 24-FEB-1998; 98FR-00002212.

XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX Cheneboux DMB, Delagneau JH, Gadelle SX, Rieunier FY;

XX WPI; 1998-583190/49.

XX New synthetic peptide(s) - useful for, e.g. detecting infection by human

XX immune deficiency virus of group O.

XX Claim 6; Page 42; 55pp; French.

XX AAW80459-74 represent synthetic peptides (either linear or cyclised by

XX Cys-Cys disulphide bonds). The peptides represent variable sequences of

XX connected around short highly conserved sequences present in isolates of

XX group O human immune deficiency virus (HIV). The peptides are useful as

XX immunological reagents for detecting infection by group O human immune

XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)

XX (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 22 AA;

XX Query Match 100.0%; Score 127; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 9.6e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNET 22  
DB 1 LLSSWGCKGRLVCYTSVQVNET 22

RESULT 2  
AAW80466  
ID AAW80466 standard; peptide; 22 AA.  
XX AC  
XX AAW80466;  
XX DT 27-AUG-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 28-JAN-1999 (first entry)  
XX DE Peptide derived from a conserved sequence of group O human HIV.  
XX KW Group O human immune deficiency virus; HIV; detection; infection.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus 1.  
XX PN WO9845323-A1.  
XX PD 15-OCT-1998.  
XX PF AC  
XX DT 27-AUG-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 28-JAN-1999 (first entry)  
XX DE Peptide derived from a conserved sequence of group O human HIV.  
XX KW Group O human immune deficiency virus; HIV; detection; infection.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus 1.  
XX PN WO9845323-A1.  
XX PD 15-OCT-1998.  
XX PF 06-APR-1998; 98WO-FR000691.  
XX PR 09-APR-1997; 97FR-00004356.  
XX PR 24-FEB-1998; 98FR-00002212.  
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX XX  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX DR  
XX PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
XX PT immune deficiency virus of group O.  
XX PS Claim 6; Page 43; 55pp; French.  
XX XX  
XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX CC connected around short highly conserved sequences present in isolates of  
XX CC group O human immune deficiency virus (HIV). The peptides are useful as  
XX CC immunological reagents for detecting infection by group O human immune  
XX CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX CC (Updated on 27-AUG-2003 to correct OS field.)  
XX SQ Sequence 22 AA;

Query Match 97.6%; Score 124; DB 2; Length 22;  
Best Local Similarity 95.5%; Pred. No. 2.4e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNET 22  
DB 1 LLSSWGCKGRLVCYTSVQVNET 22

RESULT 3  
AAW80462  
ID AAW80462 standard; peptide; 22 AA.  
XX AC  
XX AAW80462;  
XX XX  
XX DT 27-AUG-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 28-JAN-1999 (first entry)

XX DE Peptide derived from a conserved sequence of group O human HIV.  
XX KW Group O human immune deficiency virus; HIV; detection; infection.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus 1.  
XX PN WO9845323-A1.  
XX PD 15-OCT-1998.  
XX PF 06-APR-1998; 98WO-FR000691.  
XX PR 09-APR-1997; 97FR-00004356.  
XX PR 24-FEB-1998; 98FR-00002212.  
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
XX XX  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
XX DR  
XX PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
XX PT immune deficiency virus of group O.  
XX PS Claim 6; Page 42; 55pp; French.  
XX XX  
XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
XX CC connected around short highly conserved sequences present in isolates of  
XX CC group O human immune deficiency virus (HIV). The peptides are useful as  
XX CC immunological reagents for detecting infection by group O human immune  
XX CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX CC (Updated on 27-AUG-2003 to correct OS field.)  
XX SQ Sequence 22 AA;

Query Match 96.1%; Score 122; DB 2; Length 22;  
Best Local Similarity 95.5%; Pred. No. 4.3e-09;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNET 22  
DB 1 LLSSWGCKGRLVCYTSVQVNET 22

RESULT 4  
AAW07346  
ID AAW07346 standard; peptide; 40 AA.  
XX AC AAW07346;  
XX XX  
XX DT 16-OCT-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 03-JUN-1997 (first entry)  
XX XX  
XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
XX KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
XX KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
XX KW immunogen; antibody.  
XX OS Human immunodeficiency virus 1.  
XX PN WO9627013-A1.  
XX PD 06-SEP-1996.  
XX XX  
XX DT 26-FEB-1996; 96WO-FR000294.  
XX PR 27-FEB-1995; 95FR-00002236.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;  
 PI WPI; 1996-412779/41.  
 XX N-PSDB; AAT44922.  
 DR New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 XX antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX Claim 12; Page 34; 71pp; French.  
 XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MVP180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC CCMC I-1544 (BCF02 (ES)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 40 AA;  
 Query Match 94.5%; Score 120; DB 2; Length 40;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 19 LLNSWGCKGRLVCYTSVQWNET 40  
 RESULT 5  
 ID AAW80463  
 AC AAW80463 standard; peptide; 22 AA.  
 XX AAW80463;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 Query Match 92.9%; Score 118; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 1 LLQSWGCKGRLVCYTSVQWNET 22  
 RESULT 6  
 ID AAW80460  
 AC AAW80460 standard; peptide; 22 AA.  
 XX AAW80460;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 Query Match 94.5%; Score 120; DB 2; Length 40;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 19 LLNSWGCKGRLVCYTSVQWNET 40  
 RESULT 5  
 ID AAW80463  
 AC AAW80463 standard; peptide; 22 AA.  
 XX AAW80463;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 Query Match 92.9%; Score 118; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 1 LLQSWGCKGRLVCYTSVQWNET 22  
 RESULT 6  
 ID AAW80460  
 AC AAW80460 standard; peptide; 22 AA.  
 XX AAW80460;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;

DR WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 Query Match 92.9%; Score 118; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 1 LLQSWGCKGRLVCYTSVQWNET 22  
 RESULT 6  
 ID AAW80460  
 AC AAW80460 standard; peptide; 22 AA.  
 XX AAW80460;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 Query Match 92.9%; Score 118; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNET 22  
 Db 1 LLQSWGCKGRLVCYTSVQWNET 22  
 RESULT 6  
 ID AAW80460  
 AC AAW80460 standard; peptide; 22 AA.  
 XX AAW80460;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 DT 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;

```
Query Match          92.9%; Score 118; DB 2; Length 22;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRGLVCYTSVQWNET 22
Db 1 LLSLWGCRGLVCYTSVQWNET 22

RESULT 7
AAW80469
ID AAW80469 standard; peptide; 32 AA.
XX
AC AAW80469;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
KW Group O human immune deficiency virus; HIV; detection; infection.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
PN WO9845323-A1.
PD 15-OCT-1998.
PF 06-APR-1998; 98WO-FR000691.
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
KW Group O human immune deficiency virus; HIV; detection; infection.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
PN WO9845323-A1.
PD 15-OCT-1998.
PF 06-APR-1998; 98WO-FR000691.
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX
DR WPI; 1998-583190/49.
XX
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
PS Claim 6; Page 44; 55pp; French.
XX
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 32 AA;

Query Match          92.1%; Score 117; DB 2; Length 32;
Best Local Similarity 86.4%; Pred. No. 2.8e-08;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRGLVCYTSVQWNET 22
Db 11 LLSWGCRGLVCYTSVRWNET 32

RESULT 8
AAW80465
ID AAW80465 standard; peptide; 22 AA.
XX
AC AAW80465;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
XX

Query Match          89.0%; Score 113; DB 2; Length 22;
Best Local Similarity 86.4%; Pred. No. 6.4e-08;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRGLVCYTSVQWNET 22
Db 1 LLSLWGCRGLVCYTSVQWNET 22

RESULT 9
AAW80464
ID AAW80464 standard; peptide; 22 AA.
XX
AC AAW80464;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
KW Group O human immune deficiency virus; HIV; detection; infection.
XX
OS Synthetic.
OS Human immunodeficiency virus 1.
PN WO9845323-A1.
PD 15-OCT-1998.
PF 06-APR-1998; 98WO-FR000691.
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX
DR WPI; 1998-583190/49.
XX
PT New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX
PS Claim 6; Page 43; 55pp; French.
XX
CC AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 22 AA;
```



```

CC (env): group M containing sub-groups A-G, and group O containing the
CC strains ANT70 and MPV5180. The invention relates to the discovery of
CC several new strains of HIV-1 which can be placed in group O, based on the
CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
CC and AA0707329-64). The novel strains have been deposited as retroviruses
CC CNEC I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
CC the strain BCF13 and corresponds to a fragment of the gp41 protein
CC encoded by the env gene. The nucleic acids can be used to detect gp. O
CC HIV-1 strains by hybridisation or (as primers) by gene amplification.
CC also for screening and typing of such strains. Peptides encoded by the
CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O
CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
CC 2003 to standardise OS field)
XX
XX
SQ Sequence 40 AA;

Query Match          89.0%; Score 113; DB 2; Length 40;
Best Local Similarity 86.4%; Pred. No. 1.1e-07;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0

QY      1 LLSSMGCKGRLVCYTSVQNNET 22
DB      ||:|||||:|||||:|||||
        19 LLSSMGCKGRLVCYTSVENWNT 40

RESULT 11
AAY05565
ID AAY05565 standard; protein; 113 AA.
XX AC AAY05565;
XX AC
DT 17-OCT-2003 (revised)
DT 19-JUL-1999 (first entry)
XX
XX DE HIV-1 group O isolate MP539-PBMC gp41 antigen.
XX
XX HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine;
XX diagnosis; AIDS.
XX Human immunodeficiency virus 1.
XX
XX Key Location/Qualifiers
FH Misc-difference 65
FT /note= "encoded by AMW"
FT Misc-difference 74
FT /note= "encoded by ATR"
FT Misc-difference 84
FT /note= "encoded by GAK"
FT Misc-difference 86
FT /note= "encoded by AGY"
XX
XX WO9904011-A2.
XX
XX 28-JAN-1999.
XX
XX 20-JUL-1998; 98WO-EP004522.
XX
XX 18-JUL-1997; 97EP-00870110.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;
XX
XX WPI; 1999-132255/11.
XX DR N-PSDB; AAY05565.
XX
XX New isolated HIV-1 group O strains - used to produce polynucleotides,
XX antigens and antibodies for use in diagnosis and in vaccines for
XX prevention of HIV-1 infection.
XX
XX Claim 3; Fig 6; 162pp; English.
XX
XX
XX

```

CC The present sequence is an antigen of the gp41 protein of HIV-1 group O  
 CC (Outlier) strain MP539-PBMC, a Cameroon isolate. The invention relates to  
 CC new HIV-1 group O antigens (see AAY05546-625), and the use of these  
 CC antigens, or nucleic acids encoding them (see AAX25154-80), in the  
 CC diagnosis and prophylaxis of AIDS. They can be used as reagents for  
 CC detecting HIV-1 group O infection and for differentiating different types  
 CC of HIV-1 group O infection. Vaccines that provide protective immunity  
 CC against HIV-1 infection, in particular against HIV-1 group O infection,  
 CC comprise at least one HIV-1 type O antigen, a nucleic acid encoding such  
 CC an antigen, a virus-like particle comprising such an antigen, or an  
 CC attenuated form of an HIV-1 type O strain. The invention also relates to  
 CC new HIV-1 group O strains, mostly from patients from Cameroon and its  
 CC neighbouring countries. (Updated on 17-OCT-2003 to standardise OS field)  
 CC  
 CC Sequence 113 AA;

Query Match 89.0%; Score 113; DB 2; Length 113;  
 Best Local Similarity 81.8%; Pred. No. 3e-07;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNET 22  
 ||:|||||:|||||:|||||:  
 Db 39 LLLNLWGCKGRLVCYTSVQWNET 60

RESULT 12  
 AAY77373  
 ID AAY77373 standard; protein; 200 AA.

XX AC AAY77373;

XX DT 22-MAY-2000 (first entry)

XX DE HIV-1 group O env gp120/gp41 pGO-5 recombinant protein, SEQ ID NO:56.

XX KW HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;  
 KW immunoassay; positive control; affinity purification; therapeutic;  
 KW Escherichia coli; antigen; synthetic gene construction; mutagen;  
 KW deletion mutation.

XX OS Human immunodeficiency virus 1; group O isolate HAM112.

XX OS Synthetic.

XX PN WO200004383-A2.

XX PD 27-JAN-2000.

XX PF 09-JUL-1999; 99WO-US015469.

XX PR 14-JUL-1998; 98US-00115171.

XX PA (ABBO ) ABBOTT LAB.

XX PI Scheffell JW, Hackett JR, Tyner JD, Hickman RK;

XX DR WPI; 2000-171290/15.

XX DR N-PSDB; AAZ90284.

XX FT Novel monoclonal antibodies useful as positive control reagent for  
 PT detecting human immunodeficiency virus infections and diagnosing,  
 PT evaluating or prognosing viral disease.

XX PS Example 3; Fig 9; 148pp; English.

XX CC The invention relates to anti-HIV-1 group O monoclonal antibodies, which  
 CC may be used as positive control reagents in immunoassays to detect and  
 CC differentiate HIV-1 infections. The invention also encompasses a  
 CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,  
 CC which has no more than 15% cross reactivity to a corresponding antigen  
 CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of  
 CC using a monoclonal antibody as a positive control reagent in an  
 CC immunoassay for the detection of anti HIV-1 group O antibodies. The  
 CC monoclonal antibodies are useful as positive control reagents in

CC immunoassays capable of detecting anti-HIV-1 Group O antibodies. Such  
 CC immunoassays involve coupling a monoclonal antibody with HIV group-1  
 CC antigen and detecting the antigen-antibody complex. The monoclonal  
 CC antibodies of the invention would be used to ensure that the reagents  
 CC provided to detect HIV-1 group O antibody were performing properly. The  
 CC monoclonal antibodies may also can be immobilised on a matrix and used  
 CC for affinity purification of specific HIV-1 group O-derived proteins from  
 CC cell cultures or biological tissues. The monoclonal antibodies can also  
 CC be used for generating chimeric antibodies for therapeutic use. Different  
 CC synthetic, recombinant or purified antibodies which identify different  
 CC epitopes of HIV antigens can be used in combination in assay to diagnose,  
 CC evaluate, or prognosticate HIV disease condition. The monoclonal  
 CC antibodies are also useful for differentiating HIV-1 Group O antigens  
 CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent  
 CC recombinant HIV-1 group O env antigens encoded by the synthetic genes  
 CC AAZ90280-290286. The recombinant HIV-1 env proteins contain various  
 CC deletions relative to the native HAM112 isolate env protein (AAY77376).  
 CC The recombinant HIV-1 group O antigens were purified and used to screen  
 CC hybridoma cultures  
 CC  
 CC Sequence 200 AA;

Query Match 89.0%; Score 113; DB 3; Length 200;  
 Best Local Similarity 81.8%; Pred. No. 5.2e-07;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNET 22

Db 112 LLLNLWGCKGRLVCYTSVQWNET 133

RESULT 13

AAY09499

ID AAY09499 standard; protein; 215 AA.

XX AC AAY09499;

XX DT 17-OCT-2003 (revised)

XX DT 15-JUL-1999 (first entry)

XX DE HIV-1 Group O env polypeptide pGO-8PL.

XX KW HIV; human immunodeficiency virus; antigen; detection; antibody;  
 KW differentiation; Group O; env; immunogen; immunoassay.

XX OS Human immunodeficiency virus 1.

XX PN WO9909179-A2.

XX PD 25-FEB-1999.

XX PF 17-AUG-1998; 98WO-US017014.

XX PR 15-AUG-1997; 97US-00911824.

XX PA (ABBO ) ABBOTT LAB.

XX PI Hackett JR, Yamaguchi J, Golden AM, Brennan CA, Hickman RK;

XX DR WPI; 1999-190167/16.

XX DR N-PSDB; AAX56078.

XX FT New isolated HIV-1 Group O env polypeptides - used for the detection of  
 PT anti-HIV antibodies and for the production of antibodies for use in  
 PT detection, purification and therapy.

XX PS Claim 17; Fig 5; 138pp; English.

XX CC The present invention describes (A) an isolated HIV-1 Group O env  
 CC polypeptide. Also described are: (1) an isolated HIV-1 Group O env  
 CC polypeptide comprising an immunoreactive portion of a polypeptide as in  
 CC (A); (2) a polynucleotide (PN) encoding a polypeptide as in (A) or (1);  
 CC (3) an antigen construct comprising a first HIV-1 Group O env polypeptide

CC fused to a second HIV-1 Group O env polypeptide; (4) an antigen construct  
 CC comprising a fusion of at least one HIV-1 Group O env polypeptide with at  
 CC least one HIV-1 Group M env polypeptide; (5) an antigen construct  
 CC comprising a fusion of a first HIV-1 env polypeptide, a second HIV-1 env  
 CC polypeptide, and at least one additional HIV-1 polypeptide; (6) an  
 CC antigen construct comprising a first HIV-2 env polypeptide fused to a  
 CC second HIV-2 env polypeptide; (7) a PN encoding an antigen construct as  
 CC in (3)-(6); (8) an expression vector comprising a PN as in (7); (9) a  
 CC host cell transformed by an expression vector as in (8); and (10) an  
 CC immunoassay kit for the detection of antibodies to HIV-1 comprising an  
 CC antigen construct as in (3)-(6). The antigen constructs can be used for  
 CC the detection of anti-HIV-1 antibodies in test samples. They can also be  
 CC used as immunogens to produce antibodies. The antibodies can be used to  
 CC purify HIV polypeptides, for therapy and for detection of HIV  
 CC polypeptides. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 215 AA;

Query Match 89.0%; Score 113; DB 2; Length 215;  
 Best Local Similarity 81.8%; Pred. No. 5.6e-07;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
 ||: |||||: |||||: |||||  
 Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 14  
 AAY06983  
 ID AAY06983 standard; protein; 215 AA.

XX AC AAY06983;

XX DT 06-JUL-1999 (first entry)

XX DE Recombinant pGO-8PL protein.

XX KW HIV-1; HIV-2; immobilised capture reagent; capillary action; screening;  
 XX antibody; assay.

XX OS Synthetic.

XX OS Human immunodeficiency virus 1.

XX FH Key Location/Qualifiers  
 FT Protein 2..46  
 FT /note= "gp120 sequence"  
 FT Peptide 47..245  
 FT /note= "gp41 sequence"

XX WO9909410-A2.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-US016506.

XX PR 15-AUG-1997; 97US-00912129.

XX PA (ABBO ) ABBOTT LAB.

XX PI Vallari AS, Hackett JR, Hickman RK, Varitek V, Necklaws EC;  
 XX Golden AM, Brennan CA, Devare SG;

XX DR WPI; 1999-190224/16.

XX DR N-PSDB; AAX37193.

XX PT New rapid assay for antibodies to HIV-1 groups O and M, and HIV-2 - can  
 XX be used in field assay, requiring no electricity and less specialised  
 XX equipment.

XX PS Claim 1; Fig 5; 104pp; English.

XX CC The invention relates to a rapid assay for simultaneous detection and  
 CC differentiation of antibodies to HIV-1 groups O and M, and HIV-2. The

CC method comprises (a) contacting the sample with a strip containing at  
 CC least one immobilised capture reagent per analyte and on which the sample  
 CC moves from the proximal to the distal end by capillary action, under  
 CC conditions sufficient to form capture reagent/analyte complexes, and (b)  
 CC determining the presence of analyte(s) by detecting a visible colour  
 CC change at the capture reagent site on the strip wherein the capture  
 CC reagent for HIV-1 group O comprises a polypeptide shown in AAY06977-80  
 CC and AAY06983-84; and that for HIV-1 group M comprises a polypeptide shown  
 CC in AAY06982; and that for HIV-2 comprises the polypeptide shown in  
 CC AAY06981. The invention is used to screen patients for antibodies to HIV-  
 CC 1 types O and M, and HIV-2. The invention will be particularly useful in  
 CC places and situation where equipment and/or electricity is not available.  
 CC The invention provides a screening method which is faster and requires a  
 CC less equipment than prior art methods. The present sequence represents a  
 CC amino acid sequence of the recombinant pGO-8PL protein which acts as a  
 CC capture reagent for HIV-1 group O  
 XX  
 SQ Sequence 215 AA;

Query Match 89.0%; Score 113; DB 2; Length 215;  
 Best Local Similarity 81.8%; Pred. No. 5.6e-07;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
 ||: |||||: |||||: |||||  
 Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 15  
 AAY77374  
 ID AAY77374 standard; protein; 215 AA.

XX AC AAY77374;

XX DT 22-MAY-2000 (first entry)

XX DE HIV-1 group O env gp120/gp41 pGO-8PL recombinant protein, SEQ ID NO:58.

XX KW HIV-1 group O; env; gp120; gp41; glycoprotein; monoclonal antibody;  
 XX immunoassay; positive control; affinity purification; therapeutic;  
 XX Escherichia coli; antigen; synthetic gene construction; mutagen;  
 XX deletion mutation.

XX OS Human immunodeficiency virus 1; group O isolate HAM112.

XX OS Synthetic.

XX PN WO200004383-A2.

XX PD 27-JAN-2000.

XX PF 09-JUL-1999; 99WO-US015469.

XX PR 14-JUL-1998; 98US-00115171.

XX PA (ABBO ) ABBOTT LAB.

XX PI Scheffel JW, Hackett JR, Tyner JD, Hickman RK;

XX DR WPI; 2000-171290/15.

XX DR N-PSDB; AAZ90285.

XX PT Novel monoclonal antibodies useful as positive control reagent for  
 XX detecting human immunodeficiency virus infections and diagnosing,  
 XX evaluating or prognosing viral disease.

XX PS Example 3; Fig 5; 148pp; English.

XX CC The invention relates to anti-HIV-1 group O monoclonal antibodies, which  
 CC may be used as positive control reagents in immunoassays to detect and  
 CC differentiate HIV-1 infections. The invention also encompasses a  
 CC monoclonal antibody which binds specifically to an HIV-1 group O antigen,  
 CC which has no more than 15% cross reactivity to a corresponding antigen  
 CC selected from HIV-1 group M antigens and HIV-2 antigens; and a method of

CC using a monoclonal antibody as a positive control reagent in an  
CC immunoassay for the detection of anti HIV-1 group O antibodies. The  
CC monoclonal antibodies are useful as positive control reagents in  
CC immunoassays capable of detecting anti-HIV-1 Group O antibodies. Such  
CC immunoassays involve coupling a monoclonal antibody with HIV group-1  
CC antigen and detecting the antigen-antibody complex. The monoclonal  
CC antibodies of the invention would be used to ensure that the reagents  
CC provided to detect HIV-1 group O antibody were performing properly. The  
CC monoclonal antibodies may also can be immobilised on a matrix and used  
CC for affinity purification of specific HIV-1 group O-derived proteins from  
CC cell cultures or biological tissues. The monoclonal antibodies can also  
CC be used for generating chimeric antibodies for therapeutic use. Different  
CC synthetic, recombinant or purified antibodies which identify different  
CC epitopes of HIV antigens can be used in combination in assay to diagnose,  
CC evaluate, or prognosticate HIV disease condition. The monoclonal  
CC antibodies are also useful for differentiating HIV-1 Group O antigens  
CC from HIV-group M and HIV-2 antigens. Sequences AAY77369-Y77375 represent  
CC recombinant HIV-1 group O env antigens encoded by the synthetic genes  
CC AAZ90280-Z90286. The recombinant HIV-1 env proteins contain various  
CC deletions relative to the native HAM12 isolate env protein (AAY77376).  
CC The recombinant HIV-1 group O antigens were purified and used to screen  
CC hybridoma cultures  
XX  
SQ Sequence 215 AA;  
Query Match 89.0%; Score 113; DB 3; Length 215;  
Best Local Similarity 81.8%; Pred. No. 5.6e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LLSWGCGRGLVCYTSVQWNET 22  
||:|||||:|||||:  
Db 127 LNLWGCKGRGLCYTSVKWNET 148

Search completed: May 7, 2004, 17:42:43  
Job time : 42.8739 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 12.1092 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-3

Perfect score: 127

Sequence: 1 LLSWCKGRLVCYTSVQWNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/2/iaa/backfile1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	94.5	40	3	US-08-894-699-39
2	120	94.5	40	3	US-09-444-410-39
3	113	89.0	40	3	US-08-894-699-68
4	113	89.0	40	3	US-09-444-410-68
5	113	89.0	113	4	US-09-462-917A-40
6	113	89.0	215	2	US-08-912-129A-58
7	113	89.0	245	2	US-08-912-129A-48
8	113	89.0	373	2	US-08-912-129A-52
9	113	89.0	460	2	US-08-912-129A-60
10	113	89.0	490	2	US-08-912-129A-50
11	113	89.0	618	2	US-08-912-129A-54
12	113	89.0	715	4	US-09-462-917A-134
13	113	89.0	873	2	US-08-912-129A-61
14	111	87.4	33	3	US-09-433-428D-6
15	110	86.6	33	3	US-09-433-428D-25
16	110	86.6	40	3	US-08-894-699-36
17	110	86.6	40	3	US-09-444-410-36
18	110	86.6	113	4	US-09-462-917A-28
19	110	86.6	116	4	US-09-462-917A-20
20	110	86.6	356	1	US-08-602-713-12
21	110	86.6	356	3	US-08-989-493-12
22	110	86.6	356	4	US-09-610-271-12
23	109	85.8	37	4	US-08-817-441-86
24	109	85.8	41	3	US-08-894-699-69
25	109	85.8	41	3	US-09-444-410-69
26	109	85.8	113	4	US-09-462-917A-2
27	109	85.8	113	4	US-09-462-917A-18

28	109	85.8	117	4	US-09-462-917A-6	Sequence 6, Appl
29	109	85.8	200	3	US-08-965-056-104	Sequence 104, App
30	109	85.8	862	4	US-09-206-551-15	Sequence 15, Appl
31	108	85.0	23	4	US-09-462-917A-137	Sequence 137, App
32	108	85.0	33	4	US-09-433-428D-30	Sequence 30, Appl
33	108	85.0	41	3	US-08-894-699-67	Sequence 67, Appl
34	108	85.0	41	3	US-09-444-410-67	Sequence 67, Appl
35	108	85.0	110	4	US-09-462-917A-14	Sequence 14, Appl
36	108	85.0	115	4	US-09-462-917A-24	Sequence 24, Appl
37	108	85.0	116	4	US-09-462-917A-22	Sequence 22, Appl
38	107	84.3	23	4	US-09-462-917A-95	Sequence 95, Appl
39	107	84.3	33	4	US-09-433-428D-8	Sequence 8, Appl
40	107	84.3	40	3	US-08-894-699-41	Sequence 41, Appl
41	107	84.3	40	3	US-08-894-699-42	Sequence 42, Appl
42	107	84.3	40	3	US-09-444-410-41	Sequence 41, Appl
43	107	84.3	40	3	US-09-444-410-42	Sequence 42, Appl
44	107	84.3	110	4	US-09-462-917A-16	Sequence 16, Appl
45	107	84.3	113	4	US-09-462-917A-10	Sequence 10, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-894-699-39

Query Match 94.5%; Score 120; DB 3; Length 40;  
Best Local Similarity 90.9%; Pred. No. 1.4e-10;  
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||  
Db 19 LLNSWGCKGRVCYTSVKWNET 40

## RESULT 2

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/444,410  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/894,699  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 94.5%; Score 120; DB 3; Length 40;  
Best Local Similarity 90.9%; Pred. No. 1.4e-10;  
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||  
Db 19 LLNSWGCKGRVCYTSVKWNET 40

## RESULT 3

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-68

Query Match 89.0%; Score 113; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 1.5e-09;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||  
Db 19 LLNSWGCKGRVCYTSVEWNET 40

## RESULT 4

US-09-444-410-68  
; Sequence 68, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE







```

; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancakers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-912-129A-60

Query Match      89.0%; Score 113; DB 2; Length 460;
Best Local Similarity 81.8%; Pred. No. 1.8e-08;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 LLSSWCKGRILCYTSVQWNET 22
DB      372 LLNLWGCKGRILCYTSVQWNET 393

RESULT 10
US-08-912-129A-50
; Sequence 50, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancakers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-54
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-50

Query Match      89.0%; Score 113; DB 2; Length 490;
Best Local Similarity 81.8%; Pred. No. 1.9e-08;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 LLSSWCKGRILCYTSVQWNET 22
DB      372 LLNLWGCKGRILCYTSVQWNET 393

RESULT 11
US-08-912-129A-54
; Sequence 54, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancakers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-54
```

```
Query Match      89.0%; Score 113; DB 2; Length 618;
Best Local Similarity 81.8%; Pred. No. 2.4e-08;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 LLSWGCKGRLVCYTSVQWNET 22
Db      372 LNLWGCKGRLVCYTSVKWNET 393

RESULT 12
US-09-462-917A-134
; Sequence 134, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NPUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Human
US-09-462-917A-134

Query Match      89.0%; Score 113; DB 4; Length 715;
Best Local Similarity 86.4%; Pred. No. 2.8e-08;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 LLSWGCKGRLVCYTSVQWNET 22
Db      628 LNLWGCKGRLVCYTSVQWNET 649

RESULT 13
US-08-912-129A-61
; Sequence 61, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancckers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 873 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-912-129A-61

Query Match      89.0%; Score 113; DB 2; Length 873;
Best Local Similarity 81.8%; Pred. No. 3.4e-08;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 LLSWGCKGRLVCYTSVQWNET 22
Db      601 LNLWGCKGRLVCYTSVKWNET 622

RESULT 14
US-09-433-428D-6
; Sequence 6, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-6

Query Match      87.4%; Score 111; DB 3; Length 33;
Best Local Similarity 90.0%; Pred. No. 2.3e-09;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 LLSWGCKGRLVCYTSVQWN 20
Db      14 LLSWGCKGRLVCYTSVEWN 33

RESULT 15
US-09-433-428D-25
; Sequence 25, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
```

US-09-433-428D-25

Query Match 86.6%; Score 110; DB 3; Length 33;  
Best Local Similarity 90.0%; Pred. No. 3.2e-09;  
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLSSWGCKGRLVCYTSVQWN 20  
Db 14 LLNSWGCKGRLVCYTSVKWN 33

Search completed: May 7, 2004, 17:53:25  
Job time : 12.1092 secs

[illegible]

Result No.	Score	Match	Length	ID	Description
1	113	89.0	113	14	Sequence 40, Appl
2	113	89.0	215	8	Sequence 58, Appl
3	113	89.0	245	8	Sequence 48, Appl
4	113	89.0	281	8	Sequence 120, Appl
5	113	89.0	373	8	Sequence 52, Appl
6	113	89.0	460	8	Sequence 60, Appl
7	113	89.0	488	8	Sequence 95, Appl
8	113	89.0	490	8	Sequence 50, Appl
9	113	89.0	526	8	Sequence 97, Appl
10	113	89.0	618	8	Sequence 54, Appl
11	113	89.0	706	8	Sequence 93, Appl
12	113	89.0	715	14	Sequence 134, Appl
13	113	89.0	736	8	Sequence 91, Appl
14	113	89.0	873	8	Sequence 61, Appl
15	110	86.6	113	14	Sequence 28, Appl

```
RESULT 2
US-08-911-824-58
; Sequence 58, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL
US-08-911-824-58

Query Match      89.0%; Score 113; DB 8; Length 215;
Best Local Similarity 81.8%; Pred. No. 9.4e-08;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLCYTSVQWNET 22
Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 3
US-08-911-824-48
; Sequence 48, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL
US-08-911-824-48

Query Match      89.0%; Score 113; DB 8; Length 245;
Best Local Similarity 81.8%; Pred. No. 1.1e-07;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLCYTSVQWNET 22
Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 4
US-08-911-824-60
```

```
US-08-911-824-120
; Sequence 120, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 120
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL
US-08-911-824-120

Query Match      89.0%; Score 113; DB 8; Length 281;
Best Local Similarity 81.8%; Pred. No. 1.2e-07;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLCYTSVQWNET 22
Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 5
US-08-911-824-52
; Sequence 52, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL
US-08-911-824-52

Query Match      89.0%; Score 113; DB 8; Length 373;
Best Local Similarity 81.8%; Pred. No. 1.6e-07;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLCYTSVQWNET 22
Db 127 LLNLWGCKGRLCYTSVKWNET 148

RESULT 6
US-08-911-824-60
```

; Sequence 60, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 60  
; LENGTH: 460  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS  
US-08-911-824-60

Query Match 89.0%; Score 113; DB 8; Length 460;  
Best Local Similarity 81.8%; Pred. No. 1.9e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||:|||||:|||||  
DB 372 LLNLWGCKGRLCYTSVKWNET 393

## RESULT 7

US-08-911-824-95  
; Sequence 95, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 95  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-14PL  
US-08-911-824-95

Query Match 89.0%; Score 113; DB 8; Length 488;  
Best Local Similarity 81.8%; Pred. No. 2e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||:|||||:|||||  
DB 127 LLNLWGCKGRLCYTSVKWNET 148

## RESULT 8

US-08-911-824-50  
; Sequence 50, Application US/08911824

; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 50  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS  
US-08-911-824-50

Query Match 89.0%; Score 113; DB 8; Length 490;  
Best Local Similarity 81.8%; Pred. No. 2e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||:|||||:|||||  
DB 372 LLNLWGCKGRLCYTSVKWNET 393

## RESULT 9

US-08-911-824-97  
; Sequence 97, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 97  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-15CKS  
US-08-911-824-97

Query Match 89.0%; Score 113; DB 8; Length 526;  
Best Local Similarity 81.8%; Pred. No. 2.2e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNET 22  
||:|||||:|||||:|||||  
DB 372 LLNLWGCKGRLCYTSVKWNET 393

## RESULT 10

US-08-911-824-54  
; Sequence 54, Application US/08911824  
; Publication No. US20030004323A1





; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.01  
; CURRENT APPLICATION NUMBER: US/08/911.824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 61  
; LENGTH: 873  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: HIV-1 Group O isolate HAM112  
US-08-911-824-61

Query Match 89.0%; Score 113; DB 8; Length 873;  
Best Local Similarity 81.8%; Pred. No. 3.5e-07;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRIVCYTSVQNNET 22  
Db 601 LLNLWGCKGRICITYSVKNNET 622

RESULT 15  
US-10-320-786-28  
; Sequence 28, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320.786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-28

Query Match 86.6%; Score 110; DB 14; Length 113;  
Best Local Similarity 81.8%; Pred. No. 1.3e-07;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRIVCYTSVQNNET 22  
Db 39 LLNLWGCKGRICITYSVQNNET 60

Search completed: May 7, 2004, 18:29:20  
Job time : 30.7815 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 9.2437 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-3  
Perfect score: 127  
Sequence: 1 LLSSWGCKGRLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	85.8	863	2 A53034	gag polyprotein -
2	106	83.5	104	2 S52930	Gp41 ENV protein -
3	103	81.1	877	2 S49197	envelope protein p
4	84	66.1	854	1 VCLJST	env polyprotein pr
5	83	65.4	357	2 S21990	envelope protein g
6	83	65.4	358	2 S22002	envelope protein g
7	82	64.6	358	2 S22000	envelope protein g
8	82	64.6	358	2 S70417	envelope protein g
9	81	63.8	443	2 C41621	env polyprotein P
10	81	63.8	853	2 S54384	envelope polyprote
11	81	63.8	855	1 VCLJZR	env polyprotein pr
12	80	63.0	357	2 S22006	envelope protein g
13	80	63.0	357	2 S21994	envelope protein g
14	80	63.0	357	2 S22004	envelope protein g
15	80	63.0	357	2 S21996	envelope protein g
16	80	63.0	357	2 S21992	envelope protein g
17	80	63.0	358	2 S21998	envelope protein g
18	80	63.0	445	2 A41621	env polyprotein M
19	80	63.0	454	2 B41621	env polyprotein D
20	80	63.0	843	1 H44001	env polyprotein pr
21	80	63.0	847	2 T09448	envelope glycoprot
22	80	63.0	847	2 S13289	env protein - huma
23	80	63.0	852	1 VCLJBR	env polyprotein -
24	80	63.0	852	2 T12016	envelope glycoprot
25	80	63.0	854	2 S13288	env protein - huma
26	80	63.0	855	1 VCLJAZ	env polyprotein pr
27	80	63.0	856	1 VCLJH3	env polyprotein pr
28	80	63.0	856	1 VCLJVL	env polyprotein pr
29	80	63.0	856	1 VCLJ3W	env polyprotein pr

30	80	63.0	859	1 VCLJMN	env polyprotein pr
31	80	63.0	861	1 VCLJLV	env polyprotein pr
32	80	63.0	861	1 VCLJSC	env polyprotein pr
33	80	63.0	868	1 VCLJH4	env polyprotein -
34	79	62.2	786	2 S28084	env polyprotein -
35	78.5	61.8	855	2 A45713	Env transmembrane
36	78	61.4	856	1 A44963	env polyprotein pr
37	77	60.6	846	1 VCLJND	env polyprotein pr
38	74.5	58.7	859	1 VCLJST	env polyprotein pr
39	74.5	58.7	859	2 S24571	env protein - huma
40	74.5	58.7	885	2 S04322	env polyprotein -
41	74.5	58.7	886	2 T11555	env protein - simi
42	74	58.3	729	1 VCLJXK	env polyprotein pr
43	74	58.3	861	1 VCLJKB	env polyprotein pr
44	73	57.5	851	2 S33985	env polyprotein -
45	72.5	57.1	151	2 S30448	env protein - huma

ALIGNMENTS

RESULT 1

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborght, B.; van der Groen,

J. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African h

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A:Molecule type: genomic RNA

A:Status: preliminary

A:Residues: 1-863 <VAN>

A:Cross-references: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 85.8%; Score 109; DB 2; Length 863;

Best Local Similarity 81.8%; Pred. No. 9.5e-08;

Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22

||| |||||:|||||:|

Db 594 LLSLWGCKGRLVCYTSVKWRT 615

RESULT 2

S52930

Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C:Accession: S52930

R:Cohen, J.H.M.; Ghetard, D.; Philbert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; d

submitted to the EMBL Data Library, January 1995

A:Description: A novel HIV1-O strain illustrates the diversity of the O group.

A:Reference number: S52929

A:Accession: S52930

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-104 <COH>

A:Cross-references: EMBL:X84328; NID:g695526; PIDN:CAA59066.1; PID:g695527

C:Superfamily: type E retrovirus env polyprotein

Query Match 83.5%; Score 106; DB 2; Length 104;

Best Local Similarity 72.7%; Pred. No. 4e-08;

Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22

||| |||||:|||||:|

Db 45 LLNLWGCRGKAICYTSVQWNET 66

```
RESULT 3
S49197
envelope protein precursor - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 26-Aug-1999
C:Accession: S49197
R:Charneau, P.; Boorman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy,
submitted to the EMBL Data Library, July 1994
A:Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: defi
A:Reference number: S49197
A:Accession: S49197
A:Molecule type: DNA
A:Residues: 1-877 <CHA>
A:Cross-references: EMBL:X80020; NID:g510516; PIDN:CAA56323.1; PID:g510517
A:Experimental source: isolate VAU
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-535/Product: coat protein gp120 #status predicted <CP1>
F:536-877/Product: coat protein gp41 #status predicted <CP2>
F:698-716/Domain: transmembrane #status predicted <TM1>
F:59,88,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404,
Query Match 81.1%; Score 103; DB 2; Length 877;
Best Local Similarity 72.7%; Pred. No. 6.9e-07;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNET 22
|||:|||||:|||||:|||||:
Db 607 LLNLWGCKNRLICYTSVKWNT 628

RESULT 4
VCLJSI
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: simian immunodeficiency virus SIVcpz
A:Note: host Pan troglodytes (chimpanzee)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09990
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09990
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-854 <HUE>
A:Cross-references: EMBL:X52154; NID:g59866; PIDN:CAA36407.1; PID:g59874
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-500/Product: coat protein gp120 #status predicted <CP1>
F:501-854/Product: coat protein gp41 #status predicted <CP2>
F:501-517/Domain: transmembrane #status predicted <TM1>
F:675-693/Domain: transmembrane #status predicted <TM2>
F:805-821/Domain: transmembrane #status predicted <TM3>
F:134,140,143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,45
Query Match 66.1%; Score 84; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 0.00035;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNET 22
|||:|||||:|||||:|||||:
Db 582 ILGLWGCKNRLICYTSVKWNT 603

RESULT 5
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Variety: isolate 20
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21990; S70423
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Molecule type: DNA
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70423
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332, 'X', 334-357 <STE2>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 65.0%; Pred. No. 0.00022;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNM 20
|||:|||||:|||||:|||||:
Db 93 LLGIWGCSGRLICTTAVPWN 112

RESULT 6
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Variety: isolate 3L
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <STE1>
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, 'X', 335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:g60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 358;
Best Local Similarity 65.0%; Pred. No. 0.00023;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNM 20
|||:|||||:|||||:|||||:
Db 94 LLGIWGCSGRLICTTAVPWN 113

RESULT 7
S22000
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
```

```
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Variety: isolate 20
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21990; S70423
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Molecule type: DNA
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70423
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332, 'X', 334-357 <STE2>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 65.0%; Pred. No. 0.00022;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNM 20
|||:|||||:|||||:|||||:
Db 93 LLGIWGCSGRLICTTAVPWN 112

RESULT 6
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Variety: isolate 3L
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <STE1>
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, 'X', 335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:g60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 358;
Best Local Similarity 65.0%; Pred. No. 0.00023;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQVNM 20
|||:|||||:|||||:|||||:
Db 94 LLGIWGCSGRLICTTAVPWN 113

RESULT 7
S22000
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
```

```
F:424-443/Domain: transmembrane #status predicted <TM>
P:9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: c

Query Match      63.8%; Score 81; DB 2; Length 443;
Best Local Similarity 54.5%; Pred. No. 0.00052;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy   1 LLSSWGCKGRLCVYTSVQWNET 22
    |||||:::||:||||:
Db   332 LLGIWCGSKLICITTAVPWNSS 353

RESULT 10
S54384
envelope polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C:Accession: S54384
R:Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54384
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-853 <THE>
A:Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: polyprotein

Query Match      63.8%; Score 81; DB 2; Length 853;
Best Local Similarity 54.5%; Pred. No. 0.00093;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy   1 LLSSWGCKGRLCVYTSVQWNET 22
    |||||:::||:||||:
Db   589 LLGIWCGSKLICITTVPNSS 610

RESULT 11
VCLJZR
env polyprotein precursor - human immunodeficiency virus Zr-6
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus Zr-6
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; C
Gene 52, 71-82, 1987
A>Title: Molecular characterization of human immunodeficiency virus from Zaïre: nucleot
A:Reference number: A26192; MUID:87248097; PMID:3036660
A:Accession: D26192
A:Molecule type: DNA
A:Residues: 1-855 <SRI>
A:Cross-references: GB:X03458; GB:M16322; NID:g329398; PIDN:AAA45380.1; PID:g329403
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-855/Product: env polyprotein #status predicted <MAT>
F:250-500/Product: exterior membrane glycoprotein #status predicted <EXT>
F:501-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,40

Query Match      63.8%; Score 81; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.00093;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy   1 LLSSWGCKGRLCVYTSVQWNET 22
    |||||:::||:||||:
Db   591 LLGIWCGSKLICITTVPNSS 612

RESULT 12
```

```
S22006
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70420; S22006
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191
A:Experimental source: patient L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.0006;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWN 20
   ||||| |:| |:| |:|
Db 93 LLGIWGCSGKLICTTTVPWN 112

RESULT 13
S21994
envelope protein gp120/gp41 - human immunodeficiency virus type 1
A:Variety: isolate 27B
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21994; S70421
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
A:Accession: S21994
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70421
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140,'X',142-312,'X',314-357 <ST2>
A:Cross-references: EMBL:X61355; NID:g60179
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.0006;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWN 20
   ||||| |:| |:| |:|
Db 93 LLGIWGCSGKLICTTTVPWN 112

RESULT 14
S22004
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 4B
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22004; S70419
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
```

```
A:Accession: S22004
A:Molecule type: DNA
A:Residues: 1-357 <STB1>
A:Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292,'X',294-357 <STE2>
A:Cross-references: EMBL:X61353; NID:g60188
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.0006;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWN 20
   ||||| |:| |:| |:|
Db 93 LLGIWGCSGKLICTTTVPWN 112

RESULT 15
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70422; S21996
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43624.1; PID:g1067129
A:Experimental source: patient 27L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.0006;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWN 20
   ||||| |:| |:| |:|
Db 93 LLGIWGCSGKLICTTTVPWN 112

Search completed: May 7, 2004, 17:51:01
Job time : 9.2437 secs
```



```
FT CARBOHYD 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E49AFAD4F2D9B3 CRC64;

Query Match 66.1%; Score 84; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 3.1e-05;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCVTSVQVNET 22
: ||||| : ||||| : ||||| : ||||| :
Db 582 ILGLGCGKAVCYTTPWNSS 603

RESULT 2
ID ENV_HV122 STANDARD; PRT; 853 AA.
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RA Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV5226.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```
FT CARBOHYD 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B0B8CBAFF7008 CRC64;

Query Match 63.8%; Score 81; DB 1; Length 853;
Best Local Similarity 54.5%; Pred. No. 9e-05;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCVTSVQVNET 22
: ||||| : ||||| : ||||| : ||||| :
Db 589 LLGLGCGSKLICITTPWNSS 610

RESULT 3
ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., Ranganathan P., Peorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.
```





```
Query Match          63.0%; Score 80; DB 1; Length 843;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRILVCYTSVQWN 20
    |||||:|:|:|:|:|
Db 579 LLGWGCSGKLICTTTVPWN 598

RESULT 5
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65024; AAA5072.1; --
DR FDB; 10BE; 15-MAY-97.
DR HIV; M38428; ENV$SF162.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

Query Match          63.0%; Score 80; DB 1; Length 847;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRILVCYTSVQWN 20
    |||||:|:|:|:|:|
Db 583 LLGWGCSGKLICTTTVPWN 602

RESULT 6
ENV_HV1W2
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.E., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
at risk for AIDS.";
RL Science 232:1548-1553 (1986).
CC -!- MISCELLANEOUS; ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12507; AAB12990.1; --
DR HIV; M12507; ENV$WMJ2.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 202 BY SIMILARITY.
FT DISULFID 125 193 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 215 244 BY SIMILARITY.
FT DISULFID 225 236 BY SIMILARITY.
FT DISULFID 293 326 BY SIMILARITY.
```





```
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 96663 MW; BE7BBF8D23C9910D CRC64;

Query Match 63.0%; Score 80; DB 1; Length 852;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLYCVTSQWN 20
DB 588 LLGWCCKGRVLYCVTSQWN 607

RESULT 10
ENV_HV1MF
ID ENV_HV1MF STANDARD; PRT; 853 AA.
AC F19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
EN
GN Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonic C., Mann A.M., Meier C.,
Wasiak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
variants diminished in the ability to induce syncytium-independent
cytolysis."
RT J. Virol. 64:3792-3803 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33943; AAA44850.1; --
CC PDB; 1A1K; 16-JUN-97.
CC HIV; M33943; ENV5MFA.
CC InterPro; IPR000328; Env GP41.
CC DR InterPro; IPR000777; GP120.
CC DR Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 30
```

```
FT CHAIN 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 337B993B6F22ABA CRC64;

Query Match 63.0%; Score 80; DB 1; Length 853;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLYCVTSQWN 20
DB 590 LLGWCCKGRVLYCVTSQWN 609

RESULT 11
ENV_HV1A2
ID ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
EN
GN Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
(RARV-2).";
RT Science 227:484-492 (1985).
RL
```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

EMBL; K02007; AAB59882.1; -;  
 DR PIR; A03976; VCLJ2A2.  
 DR HIV; K02007; ENV\$SF2.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 208 BY SIMILARITY.  
 FT DISULFID 125 199 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 221 250 BY SIMILARITY.  
 FT DISULFID 231 242 BY SIMILARITY.  
 FT DISULFID 299 333 BY SIMILARITY.  
 FT DISULFID 380 442 BY SIMILARITY.  
 FT DISULFID 387 415 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 855;  
 Best Local Similarity 60.0%; Pred. No. 0.00013;  
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLYCYTSVQWN 20

Db 591 LLGIWGCGKLCICTTAVPWN 610

RESULT 12

ENV\_HVLOY  
 ID ENV\_HVLOY STANDARD; PRT; 855 AA.  
 AC P20888;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90148544; PubMed=2559749;  
 RA Huet T.; Dazza M.C.; Brun-Vezinet F.; Roelants G.E.; Wain-Hobson S.;  
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese  
 RT individual presenting an atypical western blot.";  
 RL AIDS 3:707-715(1989).  
 CC -!- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A  
 CC HEALTHY GABONESE INDIVIDUAL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M26727; AAA83397.1; -;  
 DR HIV; M26727; ENV\$OYI.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 210 BY SIMILARITY.  
 FT DISULFID 125 201 BY SIMILARITY.  
 FT DISULFID 130 162 BY SIMILARITY.  
 FT DISULFID 223 252 BY SIMILARITY.  
 FT DISULFID 233 244 BY SIMILARITY.  
 FT DISULFID 301 335 BY SIMILARITY.  
 FT DISULFID 381 442 BY SIMILARITY.  
 FT DISULFID 388 415 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 63.0%; Score 80; DB 1; Length 855;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWN 20
   |||||:|:|:|:|:|:|
Db 591 LLGWCSCGKLICTTTPWPN 610

RESULT 13
ENV HV1B1
ID _ENV HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane
DE glycoprotein (gp120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Ivanoff L., Rafalski J.A., Whitehorn E.A.,
RA Baumanister K., Dorian E.R., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.,
RA "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284 (1985).
RN [2]
RN DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells."
RL J. Biol. Chem. 265:10373-10382 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15654; AAA44205.1; --
DR PIR; A03973; VCLJH3
DR HIV; M15654; ENV$BH102.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157

```

```

FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97224 MW; 0BFPB1A18931BB27 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 856;
Best Local Similarity 60.0%; Pred. No. 0.00013;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWN 20
   |||||:|:|:|:|:|:|
Db 592 LLGWCSCGKLICTTTPWPN 611

RESULT 14
ENV HV1H2
ID _ENV HV1H2 STANDARD; PRT; 856 AA.
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus."
RL AIDS Res. Hum. Retroviruses 3:57-69 (1987).
RN [2]
RN REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; K03455; AAB50262.1; --  
DR ENBL; AF038399; AAB99976.1; --  
DR ENBL; AF038319; AAC82596.1; --  
DR PDB; 1DF4; 26-JAN-00.  
DR PDB; 1DF5; 26-JAN-00.  
DR PDB; 1DLB; 02-APR-00.  
DR PDB; 1G9W; 27-DEC-00.  
DR PDB; 1GCL; 19-AUG-98.  
DR PDB; 1GZL; 10-OCT-02.  
DR PDB; 1K33; 10-OCT-01.  
DR PDB; 1K34; 10-OCT-01.  
DR HTV; K03455; ENVSHXB2.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal;  
KW 3D-structure.  
FT SIGNAL 1 30  
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT FT DISULFID 54 74 BY SIMILARITY.  
FT FT DISULFID 119 205 BY SIMILARITY.  
FT FT DISULFID 126 196 BY SIMILARITY.  
FT FT DISULFID 131 157 BY SIMILARITY.  
FT FT DISULFID 218 247 BY SIMILARITY.  
FT FT DISULFID 228 239 BY SIMILARITY.  
FT FT DISULFID 296 331 BY SIMILARITY.  
FT FT DISULFID 378 445 BY SIMILARITY.  
FT FT DISULFID 385 418 BY SIMILARITY.  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;  
Query Match 63.0%; Score 80; DB 1; Length 856;



FT	CARBOHYD	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA; 96938 MW; 0C241332CF7E6687 CRC64;		

Query Match 63.0%; Score 80; DB 1; Length 856;  
Best Local Similarity 60.0%; Pred. No. 0.00013;  
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRVLCVTSQWN 20  
Db 592 LLGIWGCGKLICTTAVPWN 611

Search completed: May 7, 2004, 17:43:55  
Job time : 6.54622 secs

Good My friend

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 27.916 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-3

Perfect score: 127  
Sequence: 1 LLSSWGCKGRLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	94.5	216	15	Q9IEC5 human immun
2	117	92.1	116	15	Q7ZJN9 human immun
3	116	91.3	209	15	Q9IEB6 human immun
4	115	90.6	219	15	Q9IEB6 human immun
5	113	89.0	116	15	O40458 human immun
6	113	89.0	124	15	Q9IHU7 human immun
7	113	89.0	130	15	Q9IHU9 human immun
8	113	89.0	158	15	Q8J3N6 human immun
9	113	89.0	218	15	Q9IE95 human immun
10	113	89.0	230	15	Q9IEB2 human immun
11	113	89.0	536	15	Q9IEE5 human immun
12	113	89.0	538	15	Q9IED5 human immun
13	113	89.0	872	15	Q8Q7H0 human immun
14	113	89.0	879	15	Q9WIU9 human immun
15	113	89.0	880	15	Q8Q7H1 human immun
16	113	89.0	882	15	Q8Q7F9 human immun

17	113	89.0	887	15	Q8Q7H6 human immun
18	113	89.0	887	15	Q8Q7G9 human immun
19	113	89.0	900	15	Q9QN28 human immun
20	112	88.2	216	15	Q9IEC7 human immun
21	112	88.2	535	15	Q9IEE2 human immun
22	111	87.4	125	15	Q9IHU8 human immun
23	111	87.4	135	15	Q9DQL9 human immun
24	111	87.4	230	15	Q9IEA9 human immun
25	111	87.4	242	15	Q9IE31 human immun
26	111	87.4	529	15	Q9IEE2 human immun
27	110	86.6	114	15	O40448 human immun
28	110	86.6	116	15	O40459 human immun
29	110	86.6	137	15	Q9IHV5 human immun
30	110	86.6	155	15	Q8J3Q4 human immun
31	110	86.6	172	15	Q9IEB3 human immun
32	110	86.6	177	15	Q9IEB0 human immun
33	110	86.6	208	15	Q9IEA3 human immun
34	110	86.6	219	15	Q9IEC8 human immun
35	110	86.6	227	15	Q9IE99 human immun
36	110	86.6	234	15	Q9IEC2 human immun
37	110	86.6	342	15	O11942 human immun
38	110	86.6	418	15	O36547 human immun
39	110	86.6	502	15	Q9IEE3 human immun
40	110	86.6	871	15	O57074 human immun
41	110	86.6	871	15	Q8Q7I2 human immun
42	110	86.6	876	15	Q8Q7H3 human immun
43	109	85.8	114	15	O40456 human immun
44	109	85.8	114	15	O40472 human immun
45	109	85.8	118	15	O40451 human immun

ALIGNMENTS

RESULT 1

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5;  
AC Q9IEC5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF07;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236394; CAB96243.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_Gp41.  
DR Pfam; PF00517; Gp41; 1.  
DR Transmembrane.  
KW NON\_TER 1  
FT NON\_TER 216 216  
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 94.5%; Score 120; DB 15; Length 216;  
Best Local Similarity 90.9%; Pred. No. 7.5e-11;  
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWNET 22

Db 47 LLSSWGCKGRLVCYTSVQWNET 68

```
RESULT 2
Q7ZJN9
ID Q7ZJN9 PRELIMINARY; PRT; 116 AA.
AC Q7ZJN9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=0692A;
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol
RT integrase, and env gp41.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, AY214120; A061840.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match 92.1%; Score 117; DB 15; Length 116;
Best Local Similarity 86.4%; Pred. No. 1.2e-10;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|:|
Db 32 LLNSWGCKGRLVCYTSVKWNQT 53

RESULT 3
Q9IE66
ID Q9IE66 PRELIMINARY; PRT; 209 AA.
AC Q9IE66;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide, gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=BCF107;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauclore P.,
RA Deplenne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AJ238862; CAB96300.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 209 209
SQ SEQUENCE 209 AA; 24116 MW; 91AC9BE2BFE74B83 CRC64;

Query Match 91.3%; Score 116; DB 15; Length 209;
Best Local Similarity 86.4%; Pred. No. 3.1e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|:|
Db 56 LLNSWGCKGRLVCYTSVKWNQT 77
```

```
RESULT 4
Q9IEB6
ID Q9IEB6 PRELIMINARY; PRT; 219 AA.
AC Q9IEB6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=BCF14;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AJ236404; CAB96252.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 219 219
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 90.6%; Score 115; DB 15; Length 219;
Best Local Similarity 86.4%; Pred. No. 4.6e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|:|
Db 57 LLNSWGCKGRLVCYTSVKWNNT 78

RESULT 5
O40458
ID O40458 PRELIMINARY; PRT; 116 AA.
AC O40458;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope transmembrane glycoprotein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=group O;
RA Biollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;
RT "Molecular characterization of envelope transmembrane glycoprotein of
RT 14 new human immunodeficiency virus type 1 group O strains from
RT different African countries.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL, Y09774; CAA70913.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; 540F5B3B7AD849D8 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 116;
Best Local Similarity 81.8%; Pred. No. 5e-10;
```

Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLYCYTSVQWNET 22  
||:|||||:|||||:|  
Db 39 LLNLWGCKGRLYCYTSVQWNET 60

RESULT 6  
Q9IHU7 PRELIMINARY; PRT; 124 AA.  
AC Q9IHU7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97CM768;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,  
RA Plenzak D., Schable C., Lal R.B.;  
RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O."  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
DR EMBL; AF229237; AAF7191.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 15187 MW; 105D515F114450F8 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 124;  
Best Local Similarity 81.8%; Pred. No. 5.4e-10;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLYCYTSVQWNET 22  
||:|||||:|||||:|  
Db 33 LLNLWGCKGRLYCYTSVQWNET 54

RESULT 7  
Q9IHU9 PRELIMINARY; PRT; 130 AA.  
AC Q9IHU9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN GP41.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97CM798;  
RX MEDLINE=20386754; PubMed=10933623;  
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,  
RA Plenzak D., Schable C., Lal R.B.;  
RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
type 1 group O."  
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).  
DR EMBL; AF229235; AAF7191.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 130 130  
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match 89.0%; Score 113; DB 15; Length 130;  
Best Local Similarity 81.8%; Pred. No. 5.6e-10;  
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLYCYTSVQWNET 22  
||:|||||:|||||:|  
Db 40 LLNSWGCKGRLYCYTSVQWNET 61

RESULT 8  
Q8J3N6 PRELIMINARY; PRT; 158 AA.  
AC Q8J3N6;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE GP41 protein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22092513; PubMed=12097573;  
RA Boutonnet N., Janssens W., Bouton C., Verschelde J.L., Heyndrickx L.,  
RA Beirnaert E., van der Groen G., Lasters I.;  
RT "Comparison of predicted scaffold-compatible sequence variation in the  
RT triple-hairpin structure of human immunodeficiency virus type 1 gp41  
with patient data."  
RL J. Virol. 76:7595-7606(2002).  
DR EMBL; AJ428017; CAD20969.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 158 158  
SQ SEQUENCE 158 AA; 18803 MW; 6B8AD855599B3CC8 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 158;  
Best Local Similarity 81.8%; Pred. No. 6.9e-10;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLYCYTSVQWNET 22  
||:|||||:|||||:|  
Db 48 LLNLWGCKGRLYCYTSVQWNET 69

RESULT 9  
Q9IE95 PRELIMINARY; PRT; 218 AA.  
AC Q9IE95;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE GP41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YBF51;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,

```
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236425; CAB96273.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 218
SQ SEQUENCE 218 AA; 25243 MW; E7F0E1E20EF79FA8 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 218;
Best Local Similarity 81.8%; Pred. No. 9.5e-10;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
|||||:|:|:|:|:|:|
Db 54 LLSSWGCKGRLVCYTSVKWNTT 75

RESULT 10
Q9IEB2 PRELIMINARY; PRT; 230 AA.
AC Q9IEB2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF15;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236408; CAB96256.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 230
SQ SEQUENCE 230 AA; 26705 MW; C1F744F82F51302D CRC64;

Query Match 89.0%; Score 113; DB 15; Length 230;
Best Local Similarity 81.8%; Pred. No. 1e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
|||||:|:|:|:|:|:|
Db 55 LLTLWGCKGRLVCYTSVKWNET 76

RESULT 11
Q9IEE5 PRELIMINARY; PRT; 536 AA.
AC Q9IEE5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
```

```
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF15;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
Deplenne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133062; CAB96223.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 536
SQ SEQUENCE 536 AA; 60419 MW; 0F533ABAF08FE552 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 536;
Best Local Similarity 81.8%; Pred. No. 2.3e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
|||||:|:|:|:|:|:|
Db 361 LLTLWGCKGRLVCYTSVKWNET 382

RESULT 12
Q9IED5 PRELIMINARY; PRT; 538 AA.
AC Q9IED5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF51;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
Deplenne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133072; CAB96233.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 538
SQ SEQUENCE 538 AA; 60777 MW; B3C9E66A233PEF1D CRC64;

Query Match 89.0%; Score 113; DB 15; Length 538;
Best Local Similarity 81.8%; Pred. No. 2.4e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
|||||:|:|:~|:~|:~|:~|:~|
Db 374 LLSSWGCKGRLVCYTSVKWNTT 395
```

```
RESULT 13
Q8Q7H0 PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
  Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
  "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
  RT Phylogenetic Clusters.";
  AIDS Res. Hum. Retroviruses 18:269-282(2002).
RL EMBL; AF383251; AAL98873.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 524 GPI20.
FT CHAIN 132 156 V1.
FT CHAIN 164 200 V2.
FT CHAIN 302 335 V3.
FT CHAIN 395 427 V4.
FT CHAIN 458 475 V5.
FT CHAIN 525 879 GP41.
SQ SEQUENCE 872 AA; 98344 MW; 91D37F60053827E3 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 872;
Best Local Similarity 81.8%; Pred. No. 3.8e-09;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|
Db 602 LLNSWGCKGRLVCYTSVQWNET 623

RESULT 14
Q9WIU9 PRELIMINARY; PRT; 879 AA.
AC Q9WIU9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp160 precursor.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VI686;
RX MEDLINE=99223950; PubMed=10207543;
RA Janssens W., Heyndrickx L., Van der Auwera G., Nkengasong J.,
  Beirnaert E., Vereecken K., Coppens S., Willems B., Franssen K.,
  Peeters M., Ndumbe P., Delaporte E., Van der Groen G.;
  "Interpatient genetic variability of HIV-1 group O.";
  AIDS 13:41-48(1999).
RL EMBL; AF383250; AAL98872.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 524 GPI20.
FT CHAIN 132 156 V1.
FT CHAIN 164 200 V2.
FT CHAIN 302 335 V3.
FT CHAIN 395 427 V4.
FT CHAIN 458 475 V5.
FT CHAIN 525 879 GP41.
SQ SEQUENCE 879 AA; 99099 MW; BC6422ADDF1A3409 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 880;
Best Local Similarity 81.8%; Pred. No. 3.9e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|
Db 608 LLNLWGCKGRLVCYTSVQWNET 629

Search completed: May 7, 2004, 17:49:11
```

```
RL AIDS 10:903-910(1996).
DR EMBL; X96526; CAA65373.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GPI20.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Signal; Transmembrane.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 524 GPI20.
FT CHAIN 132 156 V1.
FT CHAIN 164 200 V2.
FT CHAIN 302 335 V3.
FT CHAIN 395 427 V4.
FT CHAIN 458 475 V5.
FT CHAIN 525 879 GP41.
SQ SEQUENCE 879 AA; 99099 MW; BC6422ADDF1A3409 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 879;
Best Local Similarity 81.8%; Pred. No. 3.9e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|
Db 605 LLNLWGCKGRLVCYTSVQWNET 626

RESULT 15
Q8Q7H1 PRELIMINARY; PRT; 880 AA.
AC Q8Q7H1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA407;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
  Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
  "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
  RT Phylogenetic Clusters.";
  AIDS Res. Hum. Retroviruses 18:269-282(2002).
RL EMBL; AF383250; AAL98872.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GPI20.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
SQ SEQUENCE 880 AA; 99784 MW; F50D9358BED301EE CRC64;

Query Match 89.0%; Score 113; DB 15; Length 880;
Best Local Similarity 81.8%; Pred. No. 3.9e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNET 22
||:|||||:|||||:|
Db 608 LLNLWGCKGRLVCYTSVQWNET 629

Search completed: May 7, 2004, 17:49:11
```

Job time : 27.916 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 41.8739 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-4

Perfect score: 126

Sequence: 1 LLSWGCKRLVCTYSQVMNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	22	2 AAW80462	Aaw80462 Peptide d
2	122	96.8	22	2 AAW80463	Aaw80463 Peptide d
3	122	96.8	22	2 AAW80461	Aaw80461 Peptide d
4	119	94.4	22	2 AAW80466	Aaw80466 Peptide d
5	115	91.3	40	2 AAW07346	Aaw07346 Partial s
6	113	89.7	22	2 AAW80460	Aaw80460 Peptide d
7	113	89.7	40	2 AAW07352	Aaw07352 Partial s
8	112	88.9	32	2 AAW80469	Aaw80469 Peptide d
9	112	88.9	41	2 AAW07351	Aaw07351 Partial s
10	112	88.9	113	2 AAY05565	Aay05565 HIV-1 gro
11	112	88.9	715	2 AAY05625	Aay05625 HIV-1 gro
12	111	88.1	33	3 AAB12212	Aab12212 Partial s
13	111	88.1	113	2 AAY05559	Aay05559 HIV-1 gro
14	110	87.3	33	3 AAB12231	Aab12231 Partial s
15	110	87.3	40	2 AAW07344	Aaw07344 Partial s
16	110	87.3	40	2 AAW07343	Aaw07343 Partial s
17	110	87.3	41	2 AAW07353	Aaw07353 Partial s
18	110	87.3	117	2 AAY05548	Aay05548 HIV-1 gro
19	109	86.5	116	2 AAY05555	Aay05555 HIV-1 gro
20	108	85.7	22	2 AAW80465	Aaw80465 Peptide d
21	108	85.7	22	2 AAW80464	Aaw80464 Peptide d
22	108	85.7	23	2 AAY05623	Aay05623 HIV-1 gro
23	108	85.7	33	3 AAB12236	Aab12236 Partial s
24	108	85.7	40	2 AAW07348	Aaw07348 Partial s
25	108	85.7	113	2 AAY05551	Aay05551 HIV-1 gro

26	108	85.7	113	2 AAY05550	Aay05550 HIV-1 gro
27	108	85.7	129	2 AAW69318	Aaw69318 Anti-HIV-
28	108	85.7	129	3 AAU77258	Aau77258 HIV-1 O-t
29	108	85.7	150	2 AAW69319	Aaw69319 Anti-HIV-
30	108	85.7	173	3 AAU77259	Aau77259 Protein A
31	108	85.7	200	2 AAY22908	Aay22908 SEQ ID NO
32	108	85.7	200	3 AAY77373	Aay77373 HIV-1 gro
33	108	85.7	200	5 ABG68379	Abg68379 Envelope
34	108	85.7	200	6 ABUS7786	Abu57786 Human imm
35	108	85.7	215	2 AAY09499	Aay09499 HIV-1 Gro
36	108	85.7	215	2 AAY06983	Aay06983 Recombina
37	108	85.7	215	3 AAY77374	Aay77374 HIV-1 gro
38	108	85.7	245	2 AAY09493	Aay09493 HIV-1 Gro
39	108	85.7	245	2 AAY06977	Aay06977 Recombina
40	108	85.7	245	3 AAY77369	Aay77369 HIV-1 gro
41	108	85.7	281	2 AAY09507	Aay09507 HIV-1 Gro
42	108	85.7	373	2 AAY09495	Aay09495 HIV-1 Gro
43	108	85.7	373	2 AAY06979	Aay06979 Recombina
44	108	85.7	460	2 AAY09500	Aay09500 HIV-1 Gro
45	108	85.7	460	2 AAY06984	Aay06984 Recombina

#### ALIGNMENTS

##### RESULT 1

AAW80462  
ID AAW80462 standard; peptide; 22 AA.

AC AAW80462;

XX XX

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)

XX XX

DE Peptide derived from a conserved sequence of group O human HIV.

XX XX

KW Group O human immune deficiency virus; HIV; detection; infection.

XX XX

OS Synthetic.

OS Human immunodeficiency virus 1.

XX XX

PN W09845323-AL.

XX XX

PD 15-OCT-1998.

XX XX

PF 06-APR-1998; 98WO-FR0000691.

XX XX

PR 09-APR-1997; 97FR-00004356.

XX XX

PR 24-FEB-1998; 98FR-00002212.

(SNFI ) PASTEUR SANOFI DIAGNOSTICS.

PA PA

PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;

XX XX

WPI; 1998-583190/49.

DR DR

PT New synthetic peptide(s) - useful for, e.g. detecting infection by human immune deficiency virus of group O.

XX XX

PS Claim 6; Page 42; 55pp; French.

XX XX

CC AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences connected around short highly conserved sequences present in isolates of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)

XX XX

SQ Sequence 22 AA;

Query Match 100.0%; Score 126; DB 2; Length 22;

```
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
Db 1 LLSSWGCKGRLVCYTSVQWNST 22

RESULT 2
AAW80463
ID AAW80463 standard; peptide; 22 AA.
XX AC AAW80463;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
XX Synthetic.
OS Human immunodeficiency virus 1.
XX WO9845323-A1.
XX 15-OCT-1998.
XX 06-APR-1998; 98WO-FR000691.
XX 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX Claim 6; Page 42; 55pp; French.
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 22 AA;

Query Match 96.8%; Score 122; DB 2; Length 22;
Best Local Similarity 95.5%; Pred. No. 3.6e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
Db 1 LLSSWGCKGRLVCYTSVQWNST 22

RESULT 4
AAW80466
ID AAW80466 standard; peptide; 22 AA.
XX AC AAW80466;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX Peptide derived from a conserved sequence of group O human HIV.
XX Group O human immune deficiency virus; HIV; detection; infection.
XX Synthetic.
OS Human immunodeficiency virus 1.
XX WO9845323-A1.
XX 15-OCT-1998.
XX 06-APR-1998; 98WO-FR000691.
XX 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX WPI; 1998-583190/49.
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
PT immune deficiency virus of group O.
XX Claim 6; Page 42; 55pp; French.
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
CC Cys-Cys disulphide bonds). The peptides represent variable sequences
CC connected around short highly conserved sequences present in isolates of
CC group O human immune deficiency virus (HIV). The peptides are useful as
CC immunological reagents for detecting infection by group O human immune
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 22 AA;

Query Match 96.8%; Score 122; DB 2; Length 22;
Best Local Similarity 95.5%; Pred. No. 3.6e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
Db 1 LLSSWGCKGRLVCYTSVQWNST 22

RESULT 3
AAW80461
ID AAW80461 standard; peptide; 22 AA.
XX AC AAW80461;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
```

PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 PT  
 XX Claim 6; Page 43; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 XX Sequence 22 AA;  
 SQ

Query Match 94.4%; Score 119; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 8.9e-09;  
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNST 22  
 DB 1 LLSSWGCKGRLVCYTSVQWNST 22

RESULT 5  
 AAW07346  
 ID AAW07346 standard; peptide; 40 AA.  
 XX  
 AC AAW07346;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
 XX  
 XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95PR-00002236.  
 XX  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;  
 XX  
 DR WPI; 1996-412779/41.  
 DR N-PSDB; AAT44922.  
 XX  
 XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 XX Claim 12; Page 34; 71pp; French.  
 PS  
 XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene

(env): group M containing sub-groups A-G, and group O containing the  
 strains ANT70 and MVP5180. The invention relates to the discovery of  
 several new strains of HIV-1 which can be placed in group O, based on the  
 partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC NCIM I-1544 (BCF02 (SSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification.  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 40 AA;  
 Query Match 91.3%; Score 115; DB 2; Length 40;  
 Best Local Similarity 86.4%; Pred. No. 5.2e-08;  
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTSVQWNST 22  
 DB 19 LLSSWGCKGRLVCYTSVQWNST 40

RESULT 6  
 AAW80460  
 ID AAW80460 standard; peptide; 22 AA.  
 XX  
 AC AAW80460;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX  
 KW Group O human immune deficiency virus; HIV; detection; infection.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97PR-00004356.  
 PR 24-FEB-1998; 98PR-00002212.  
 XX  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 PA  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 PT  
 XX Claim 6; Page 42; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 XX Sequence 22 AA;  
 SQ

2

FT Misc-difference 65

```

XX PF 20-JUL-1998; 98WO-EP004522.
XX PR 18-JUL-1997; 97EP-00870110.
XX PA (INNO-) INNOGENETICS NV.
XX PI Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;
XX XX WPI; 1999-132255/11.
DR N-PSDB; AAX25180.
XX XX New isolated HIV-1 group O strains - used to produce polynucleotides,
PT antigens and antibodies for use in diagnosis and in vaccines for
PT prevention of HIV-1 infection.
XX XX Claim 3; Fig 8A; 162pp; English.
XX XX The present sequence is a partial Env polypeptide of HIV-1 group O
CC (Outlier) virus isolate MP645, as deduced from part of the genome of
CC MP645 (see AAX25180). The invention relates to new HIV-1 group O
CC antigens, especially envelope protein antigens (see AAY0546-625), and
CC the use of these antigens, or nucleic acids encoding them (see AAX25154-
CC 80), in the diagnosis and prophylaxis of AIDS. They can be used as
CC reagents for detecting HIV-1 group O infection and for differentiating
CC different types of HIV-1 group O infection. Vaccines that provide
CC protective immunity against HIV-1 infection, in particular against HIV-1
CC group O infection, comprise at least one HIV-1 type O antigen, a nucleic
CC acid encoding such an antigen, a virus-like particle comprising such an
CC antigen, or an attenuated form of an HIV-1 type O strain. The invention
CC also relates to new HIV-1 group O strains, obtained from patients from
CC Cameroon, Gabon, Tchad, Nigeria, Senegal and Niger. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX XX Sequence 715 AA;

Query Match 88.9%; Score 112; DB 2; Length 715;
Best Local Similarity 86.4%; Pred. No. 1.9e-06;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQNST 22
Db 628 LLNLWGCKGRVCYTSVQNWRT 649

RESULT 12
AAB12212
ID AAB12212 standard; peptide; 33 AA.
XX AC AAB12212;
XX XX 12-SEP-2003 (revised)
DT 10-NOV-2000 (first entry)
XX XX Partial sequence of HIV-1 strain BCF13 gp41 immunodominant region.
DE XX HIV-1; AIDS, human immunodeficiency virus type 1; immunodominant region;
KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF13.
XX XX Human immunodeficiency virus 1.
OS EP1013766-A2.
XX PN 28-JUN-2000.
XX PD 29-NOV-1999; 99EP-00309491.
XX PF 30-NOV-1998; 98US-0110292P.
XX PR 08-FEB-1999; 99US-0119138P.
XX PR 04-NOV-1999; 99US-00433428.
XX XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX XX

PI De Leys R, Zheng J;
XX WPI; 2000-402205/35.
XX PT New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies.
XX XX Example 1; Fig 1; 52pp; English.
XX XX The present sequence is a partial gp41 protein of Human Immunodeficiency
CC Virus Type 1 (HIV-1) strain BCF13. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. BCF13 is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment, which
CC in turn was used to derive a consensus sequence peptide: peptide 147
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)
XX XX Sequence 33 AA;

Query Match 88.1%; Score 111; DB 3; Length 33;
Best Local Similarity 90.0%; Pred. No. 1.5e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWN 20
Db 14 LLNLWGCKGRVCYTSVEWN 33

RESULT 13
AAY05559
ID AAY05559 standard; protein; 113 AA.
XX AC AAY05559;
XX XX 17-OCT-2003 (revised)
DT 19-JUL-1999 (first entry)
XX XX HIV-1 group O isolate BSI89 gp41 antigen.
DE XX HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine;
KW diagnosis; AIDS.
XX OS Human immunodeficiency virus 1.
XX PN WO9904011-A2.
XX PD 28-JAN-1999.
XX XX 20-JUL-1998; 98WO-EP004522.
PF 18-JUL-1997; 97EP-00870110.
PR (INNO-) INNOGENETICS NV.
PA Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;
PI WPI; 1999-132255/11.
DR N-PSDB; AAX25167.
XX XX New isolated HIV-1 group O strains - used to produce polynucleotides,
PT antigens and antibodies for use in diagnosis and in vaccines for
PT prevention of HIV-1 infection.
XX XX Claim 3; Fig 6; 162pp; English.
XX XX The present sequence is an antigen of the gp41 protein of HIV-1 group O
CC (Outlier) strain BSI89, a Cameroon isolate. The invention relates to new
CC HIV-1 group O antigens (see AAY0546-625), and the use of these antigens,
CC or nucleic acids encoding them (see AAX25154-80), in the diagnosis and
CC prophylaxis of AIDS. They can be used as reagents for detecting HIV-1

```

CC group O infection and for differentiating different types of HIV-1 group  
 CC O infection. Vaccines that provide protective immunity against HIV-1  
 CC infection, in particular against HIV-1 group O infection, comprise at  
 CC least one HIV-1 type O antigen, a nucleic acid encoding such an antigen,  
 CC a virus-like particle comprising such an antigen, or an attenuated form  
 CC of an HIV-1 type O strain. The invention also relates to new HIV-1 group  
 CC O strains, mostly from patients from Cameroon and its neighbouring  
 CC countries. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 113 AA;

Query Match 88.1%; Score 111; DB 2; Length 113;  
 Best Local Similarity 81.8%; Pred. No. 4.6e-07;  
 Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWNST 22  
 ||:|||||:|||||:  
 Db 39 LLNLWGCKGRLVCYTSVQWNMT 60

RESULT 14  
 AAB12231  
 ID AAB12231 standard; peptide; 33 AA.

AC AAB12231;

XX 12-SEP-2003 (revised)  
 DT 10-NOV-2000 (first entry)

XX Partial sequence of HIV-1 strain MAN gp41 immunodominant region.

XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; MAN.

XX Human immunodeficiency virus 1.

XX EP1013766-A2.

XX 28-JUN-2000.

XX 29-NOV-1999; 99EP-00309491.

XX 30-NOV-1998; 98US-0110292P.

PR 08-FEB-1999; 99US-0119138P.

PR 04-NOV-1999; 99US-00433428.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX De Leys R, Zheng J;

XX WPI; 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.

XX Example 1; Fig 1; 52pp; English.

XX The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain MAN. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. MAN is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 33 AA;

Query Match 87.3%; Score 110; DB 3; Length 33;  
 Best Local Similarity 90.0%; Pred. No. 2e-07;  
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWN 20  
 ||:|||||:|||||:  
 Db 14 LLNSWGCKGRLVCYTSVKWN 33

RESULT 15  
 AAW07344

ID AAW07344 standard; peptide; 40 AA.

XX AAW07344;

XX 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF01 (FAN).

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.

XX Human immunodeficiency virus 1.

XX WO9627013-A1.

XX 06-SEP-1996.

XX 26-FEB-1996; 96WO-FR000294.

XX 27-FEB-1995; 95FR-00002236.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX N-PSDB; AAT44923.

XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.

XX Claim 12; Page 33; 71pp; French.

XX Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MVP5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC CINCIM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF01 (FAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification,  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)

XX SQ Sequence 40 AA;

Query Match 87.3%; Score 110; DB 2; Length 40;  
 Best Local Similarity 81.8%; Pred. No. 2.4e-07;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWNST 22  
 ||:|||||:|||||:  
 Db 19 LLNLWGCKGRLVCYTSVKWNST 40

Search completed: May 7, 2004, 17:42:43  
Job time : 41.8739 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 12.1092 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-4

Perfect score: 126

Sequence: 1 LLSWCKGRLVCYTSVQWNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A COMB.pcp.\*

2: /cgn2\_6/prodata/2/iaa/5B COMB.pcp.\*

3: /cgn2\_6/prodata/2/iaa/6A COMB.pcp.\*

4: /cgn2\_6/prodata/2/iaa/6B COMB.pcp.\*

5: /cgn2\_6/prodata/2/iaa/PCUTS COMB.pcp.\*

6: /cgn2\_6/prodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	91.3	40	3	US-08-894-699-39
2	115	91.3	40	3	US-09-444-410-39
3	113	89.7	40	3	US-08-894-699-68
4	113	89.7	40	3	US-09-444-410-68
5	112	88.9	41	3	US-08-894-699-67
6	112	88.9	41	3	US-09-444-410-67
7	112	88.9	113	4	US-09-462-917A-40
8	112	88.9	715	4	US-09-462-917A-134
9	111	88.1	33	3	US-09-433-428D-6
10	111	88.1	113	4	US-09-462-917A-28
11	110	87.3	33	3	US-09-433-428D-25
12	110	87.3	40	3	US-08-894-699-36
13	110	87.3	40	3	US-08-894-699-37
14	110	87.3	40	3	US-09-444-410-36
15	110	87.3	40	3	US-09-444-410-37
16	110	87.3	41	3	US-08-894-699-69
17	110	87.3	41	3	US-09-444-410-69
18	110	87.3	117	4	US-09-462-917A-6
19	109	86.5	116	4	US-09-462-917A-20
20	108	85.7	23	4	US-09-462-917A-137
21	108	85.7	33	3	US-09-433-428D-30
22	108	85.7	37	4	US-08-817-441-86
23	108	85.7	40	3	US-08-894-699-41
24	108	85.7	40	3	US-09-444-410-41
25	108	85.7	113	4	US-09-462-917A-10
26	108	85.7	113	4	US-09-462-917A-12
27	108	85.7	200	3	US-08-965-056-104

28	108	85.7	215	2	US-08-912-129A-58	Sequence 58, Appl
29	108	85.7	245	2	US-08-912-129A-48	Sequence 48, Appl
30	108	85.7	373	2	US-08-912-129A-52	Sequence 52, Appl
31	108	85.7	460	2	US-08-912-129A-60	Sequence 60, Appl
32	108	85.7	490	2	US-08-912-129A-50	Sequence 50, Appl
33	108	85.7	618	2	US-08-912-129A-54	Sequence 54, Appl
34	108	85.7	862	4	US-09-206-551-15	Sequence 15, Appl
35	108	85.7	873	2	US-08-912-129A-61	Sequence 61, Appl
36	107	84.9	23	4	US-09-462-917A-95	Sequence 95, Appl
37	107	84.9	33	3	US-09-433-428D-8	Sequence 8, Appl
38	107	84.9	40	3	US-08-894-699-40	Sequence 40, Appl
39	107	84.9	40	3	US-09-444-410-40	Sequence 40, Appl
40	107	84.9	110	4	US-09-462-917A-14	Sequence 14, Appl
41	107	84.9	113	4	US-09-462-917A-2	Sequence 2, Appl
42	107	84.9	113	4	US-09-462-917A-18	Sequence 18, Appl
43	107	84.9	115	4	US-09-462-917A-24	Sequence 24, Appl
44	106	84.1	110	4	US-09-462-917A-16	Sequence 16, Appl
45	105	83.3	24	4	US-09-462-917A-138	Sequence 138, App

ALIGNMENTS

RESULT 1  
US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-894-699-39

Query Match 91.3%; Score 115; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 7.8e-10;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTSVQNST 22  
||:|||||:|||||:|||||  
Db 19 LLNSWGCKGRLVCYTSVKWNET 40

RESULT 2

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-09-444-410-39

Query Match 91.3%; Score 115; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 7.8e-10;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTSVQNST 22  
||:|||||:|||||:|||||  
Db 19 LLNSWGCKGRLVCYTSVKWNET 40

RESULT 3

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,699  
FILING DATE: 01-DEC-1997

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR96/00294  
FILING DATE: 26-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-894-699-68

Query Match 89.7%; Score 113; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 1.5e-09;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLNSWGCKGRLVCYTSVQNST 22  
||:|||||:|||||:|||||  
Db 19 LLNSWGCKGRLVCYTSVQNST 40

RESULT 4

US-09-444-410-68  
; Sequence 68, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE

;; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/444,410  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/894,699  
;; FILING DATE:  
;; INFORMATION FOR SEQ ID NO: 68:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-09-444-410-68

Query Match 89.7%; Score 113; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 1.5e-09;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTSVQWNST 22  
Db 19 LLSWGCKGRLVCYTSVQWNST 40

RESULT 5  
US-08-894-699-67  
; Sequence 67, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA

;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/894,699  
;; FILING DATE: 01-DEC-1997  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR96/00294  
;; FILING DATE: 26-FEB-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614  
;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 67:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 41 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-894-699-67

Query Match 88.9%; Score 112; DB 3; Length 41;  
Best Local Similarity 81.8%; Pred. No. 2.2e-09;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRLVCYTSVQWNST 22  
Db 19 LLSWGCKGRLVCYTSVQWNST 40

RESULT 6  
US-09-444-410-67  
; Sequence 67, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/444,410  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/894,699  
; FILING DATE:  
; PRIOR APPLICATION DATA: FR 95/02236  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 41 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-444-410-67

Query Match 88.9%; Score 112; DB 3; Length 41;  
Best Local Similarity 81.8%; Pred. No. 2.2e-09;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22  
||:|||||:|||||:|||||  
DB 19 LNLWGCKGRLVCYTSVRKWNST 40

RESULT 7  
US-09-462-917A-40  
; Sequence 40, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 40  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc.feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-09-462-917A-40

Query Match 88.9%; Score 112; DB 4; Length 113;  
Best Local Similarity 81.8%; Pred. No. 6e-09;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22  
||:|||||:|||||:|||||  
DB 39 LNLWGCKGRLVCYTSVQWNKT 60

RESULT 8  
US-09-462-917A-134  
; Sequence 134, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine

; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 134  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Human  
US-09-462-917A-134

Query Match 88.9%; Score 112; DB 4; Length 715;  
Best Local Similarity 86.4%; Pred. No. 3.9e-08;  
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22  
||:|||||:|||||:|||||  
DB 628 LNLWGCKGRLVCYTSVQWNFT 649

RESULT 9  
US-09-433-428D-6  
; Sequence 6, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-6

Query Match 88.1%; Score 111; DB 3; Length 33;  
Best Local Similarity 90.0%; Pred. No. 2.4e-09;  
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWN 20  
||:|||||:|||||:|||||  
DB 14 LLSWGCKGRLVCYTSVEMN 33

RESULT 10  
US-09-462-917A-28  
; Sequence 28, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 113  
; TYPE: PRT

```

1 FILING DATE: 26-FEB-1996
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: FR 95/02236
4 FILING DATE: 27-FEB-1995
5 ATTORNEY/AGENT INFORMATION:
6 NAME: OBLON, NORMAN F.
7 REGISTRATION NUMBER: 24,614
8 REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 703-413-3000
11 TELEFAX: 703-413-2220
12 INFORMATION FOR SEQ ID NO: 36:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 40 amino acids
15 TYPE: amino acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: peptide
19 US-08-894-699-36
20
21 Query Match 87.3%; Score 110; DB 3; Length 40;
22 Best Local Similarity 81.8%; Pred. No. 4.1e-09;
23 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
24
25 QY 1 LLNSWGCKRLVCYTSVQWNST 22
26 ||:|||||:||||:|
27 Db 19 LLNSWGCKGRIVCYTSVKWNWT 40
28
29 RESULT 13
30 US-08-894-699-37
31 Sequence 37, Application US/08894699
32 Patent No. 6030769
33 GENERAL INFORMATION:
34 APPLICANT: SIMON, FRANCOIS
35 APPLICANT: SARAGOSTI, SENTOB
36 APPLICANT: LOUSSET-AJAKA, IBITISSAM
37 APPLICANT: LY, THOAI-DUONG
38 APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
39 TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
40 TITLE OF INVENTION: VIRUSES, AND USES THEREOF
41 NUMBER OF SEQUENCES: 81
42 CORRESPONDENCE ADDRESS:
43 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
44 ADDRESSES: P.C.
45 STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
46 CITY: ARLINGTON
47 STATE: VA
48 COUNTRY: USA
49 ZIP: 22202
50
51 COMPUTER READABLE FORM:
52 MEDIUM TYPE: Floppy disk
53 COMPUTER: IBM PC compatible
54 OPERATING SYSTEM: PC-DOS/MS-DOS
55 SOFTWARE: Patent In Release #1.0, Version #1.30
56 CURRENT APPLICATION DATA:
57 APPLICATION NUMBER: US/08/894,699
58 FILING DATE: 01-DEC-1997
59 CLASSIFICATION:
60 PRIOR APPLICATION DATA:
61 APPLICATION NUMBER: PCT/FR96/00294
62 FILING DATE: 26-FEB-1996
63 PRIOR APPLICATION DATA:
64 APPLICATION NUMBER: FR 95/02236
65 FILING DATE: 27-FEB-1995
66 ATTORNEY/AGENT INFORMATION:
67 NAME: OBLON, NORMAN F.
68 REGISTRATION NUMBER: 24,614
69 REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
70 TELECOMMUNICATION INFORMATION:
71 TELEPHONE: 703-413-3000
72 TELEFAX: 703-413-2220

```

; INFORMATION FOR SEQ ID NO: 37:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 40 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-894-699-37

Query Match 87.3%; Score 110; DB 3; Length 40;  
 Best Local Similarity 81.8%; Pred. No. 4.1e-09;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22  
 ||:|||||:|||||:|||||  
 Db 19 LNLWGCKGRLVCYTSVKWNST 40

RESULT 14  
 US-09-444-410-36  
 ; Sequence 36, Application US/09444410  
 ; Patent No. 6270975  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SIMON, FRANCOIS  
 ; APPLICANT: SARAGOSTI, SENTOB  
 ; APPLICANT: LOUSERT-AJAKA, IBITISSAM  
 ; APPLICANT: LY, THOAI-DUONG  
 ; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
 ; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
 ; NUMBER OF SEQUENCES: 81  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
 ; STREET: FLOOR  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/444,410  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/894,699  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 95/02236  
 ; FILING DATE: 27-FEB-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,614  
 ; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 36:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 40 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-09-444-410-36

Query Match 87.3%; Score 110; DB 3; Length 40;  
 Best Local Similarity 81.8%; Pred. No. 4.1e-09;

Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LLSWGCKGRLVCYTSVQWNST 22  
 ||:|||||:|||||:|||||  
 Db 19 LLSWGCKGRLVCYTSVKWNT 40

RESULT 15  
 US-09-444-410-37  
 ; Sequence 37, Application US/09444410  
 ; Patent No. 6270975  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SIMON, FRANCOIS  
 ; APPLICANT: SARAGOSTI, SENTOB  
 ; APPLICANT: LOUSERT-AJAKA, IBITISSAM  
 ; APPLICANT: LY, THOAI-DUONG  
 ; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
 ; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
 ; NUMBER OF SEQUENCES: 81  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
 ; STREET: FLOOR  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/444,410  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/894,699  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 95/02236  
 ; FILING DATE: 27-FEB-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,614  
 ; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 37:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 40 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-09-444-410-37

Query Match 87.3%; Score 110; DB 3; Length 40;  
 Best Local Similarity 81.8%; Pred. No. 4.1e-09;  
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22  
 ||:|||||:|||||:|||||  
 Db 19 LNLWGCKGRLVCYTSVKWNST 40

Search completed: May 7, 2004, 17:53:25  
 Job time : 12.1092 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 30.7815 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-4

Perfect score: 126

Sequence: 1 LLSSWCKGRLVCYTSVQWNST 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	88.9	113	14	US-10-320-786-40
2	112	88.9	715	14	US-10-320-786-134
3	111	88.1	113	14	US-10-320-786-28
4	110	87.3	117	14	US-10-320-786-6
5	109	86.5	116	14	US-10-320-786-20
6	108	85.7	23	14	US-10-320-786-137
7	108	85.7	37	14	US-10-026-741-86
8	108	85.7	113	14	US-10-320-786-10
9	108	85.7	113	14	US-10-320-786-12
10	108	85.7	200	9	US-09-854-816-104
11	108	85.7	215	8	US-08-911-824-58
12	108	85.7	245	8	US-08-911-824-48
13	108	85.7	281	8	US-08-911-824-120
14	108	85.7	373	8	US-08-911-824-52
15	108	85.7	460	8	US-08-911-824-60

16	108	85.7	488	8	US-08-911-824-95	Sequence 95, Appl
17	108	85.7	490	8	US-08-911-824-50	Sequence 50, Appl
18	108	85.7	526	8	US-08-911-824-97	Sequence 97, Appl
19	108	85.7	618	8	US-08-911-824-54	Sequence 54, Appl
20	108	85.7	706	8	US-08-911-824-93	Sequence 93, Appl
21	108	85.7	736	8	US-08-911-824-91	Sequence 91, Appl
22	108	85.7	862	14	US-10-369-294-15	Sequence 15, Appl
23	108	85.7	873	8	US-08-911-824-61	Sequence 61, Appl
24	107	84.9	23	14	US-10-320-786-95	Sequence 95, Appl
25	107	84.9	110	14	US-10-320-786-14	Sequence 14, Appl
26	107	84.9	113	14	US-10-320-786-2	Sequence 2, Appl
27	107	84.9	113	14	US-10-320-786-18	Sequence 18, Appl
28	107	84.9	115	14	US-10-320-786-24	Sequence 24, Appl
29	106	84.1	110	14	US-10-320-786-16	Sequence 16, Appl
30	105	83.3	24	14	US-10-320-786-138	Sequence 138, Appl
31	105	83.3	110	14	US-10-320-786-38	Sequence 38, Appl
32	105	83.3	356	14	US-10-357-400-12	Sequence 12, Appl
33	104	82.5	24	14	US-10-320-786-92	Sequence 92, Appl
34	103	81.7	23	14	US-10-320-786-102	Sequence 102, Appl
35	103	81.7	116	14	US-10-320-786-22	Sequence 22, Appl
36	102	81.0	23	14	US-10-320-786-91	Sequence 91, Appl
37	102	81.0	23	14	US-10-320-786-99	Sequence 99, Appl
38	102	81.0	23	14	US-10-320-786-100	Sequence 100, Appl
39	102	81.0	23	14	US-10-320-786-101	Sequence 101, Appl
40	102	81.0	35	14	US-10-026-741-101	Sequence 101, Appl
41	102	81.0	37	14	US-10-026-741-7	Sequence 7, Appl
42	102	81.0	37	14	US-10-026-741-90	Sequence 90, Appl
43	102	81.0	113	14	US-10-320-786-4	Sequence 4, Appl
44	102	81.0	351	14	US-10-026-741-47	Sequence 47, Appl
45	102	81.0	877	14	US-10-026-741-102	Sequence 102, Appl

ALIGNMENTS

RESULT 1  
US-10-320-786-40  
; Sequence 40, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014 DVUS01  
; CURRENT APPLICATION NUMBER: US/10320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc.feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-10-320-786-40

Query Match 88.9%; Score 112; DB 14; Length 113;  
Best Local Similarity 81.8%; Pred. No. 1e-07;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLSSWCKGRLVCYTSVQWNST 22

Db 39 LLNLWCKGRLVCYTSVQWNKT 60





FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: US 09/462,917  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 137  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(23)  
OTHER INFORMATION:  
US-10-320-786-137  
  
Query Match 85.7%; Score 108; DB 14; Length 23;  
Best Local Similarity 90.0%; Pred. No. 8.3e-08;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 LLSSWGCKGRLVCYTSVQWN 20  
Db 4 LLNLWGCKGRLVCYTSVQWN 23  
  
RESULT 7  
US-10-026-741-86  
Sequence 86, Application US/10026741  
Publication No. US20030049604A1  
GENERAL INFORMATION:  
APPLICANT: CHARNEAU, PIERRE  
CLAVEL, FRANCOISE  
BORMAN, ANDREW  
QUILLET, CAROLINE  
GUETARD, DENISE  
MONTAGNIER, LUC  
DONJON DE SAINT-MARTIN, JACQUELINE  
COHEN, JACQUES  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
SUBTYPE) ANTIGENS  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/026,741  
FILING DATE: 27-Dec-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/817,441  
FILING DATE: 31-AUG-1998  
APPLICATION NUMBER: PCT/FR 95/01391  
FILING DATE: 20-OCT-1995  
APPLICATION NUMBER: FR 9412554  
FILING DATE: 20-OCT-1994  
APPLICATION NUMBER: FR 9502526  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
US-10-026-741-86  
  
Query Match 85.7%; Score 108; DB 14; Length 37;  
Best Local Similarity 81.8%; Pred. No. 1.3e-07;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 LLSSWGCKGRLVCYTSVQWNST 22  
Db 14 LLNLWGCKGRLVCYTSVKWNRT 35  
  
RESULT 8  
US-10-320-786-10  
Sequence 10, Application US/10320786  
Publication No. US20030180759A1  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: US 09/462,917  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20  
NUMBER OF SEQ ID NOS: 152  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 10  
LENGTH: 113  
TYPE: PRT  
ORGANISM: Human  
US-10-320-786-10  
  
Query Match 85.7%; Score 108; DB 14; Length 113;  
Best Local Similarity 77.3%; Pred. No. 3.7e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 LLSSWGCKGRLVCYTSVQWNST 22  
Db 39 LLNLWGCKGRLVCYTSVKWNRT 60  
  
RESULT 9  
US-10-320-786-12  
Sequence 12, Application US/10320786  
Publication No. US20030180759A1  
GENERAL INFORMATION:  
APPLICANT: Delaporte, Eric  
APPLICANT: Peeters, Martine  
APPLICANT: Saman, Eric  
APPLICANT: Vanden Haesevelde, Marlen  
TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
CURRENT APPLICATION NUMBER: US/10/320,786  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: US 09/462,917  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: PCT/EP98/04522  
PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-12

Query Match 85.7%; Score 108; DB 14; Length 113;  
Best Local Similarity 77.3%; Pred. No. 3.7e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTSVQWNST 22  
||:|||||:|||||:|||||:  
Db 39 LLNLWGCKGRVLCYTSVKWNST 60

## RESULT 10

US-09-854-816-104  
; Sequence 104, Application US/09854816  
; Patent No. US20020151473A1

## GENERAL INFORMATION:

; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell  
; J. Christopher Phelan  
; Melissa A. Starovassnik  
; James A. Wells

; TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

; NUMBER OF SEQUENCES: 113

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA: US/09/854,816

; FILING DATE: 15-May-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/965,056

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Torchia, PhD., Timothy E.

; REGISTRATION NUMBER: 36,700

; REFERENCE/DOCKET NUMBER: P1005R2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-8674

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 104:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 200 amino acids

; TYPE: Amino Acid

; TOPOLOGY: Linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 104:

US-09-854-816-104

Query Match 85.7%; Score 108; DB 9; Length 200;  
Best Local Similarity 81.8%; Pred. No. 6.2e-07;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTSVQWNST 22  
||:|||||:|||||:|||||:  
Db 126 LLNLWGCKGRVLCYTSVKWNRT 147

## RESULT 11

US-08-911-824-58

; Sequence 58, Application US/08911824

; Publication No. US20030004323A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Hackett, John R., Jr.

; APPLICANT: Yamaguchi, Julie

; APPLICANT: Golden, Alan M.

; APPLICANT: Brennan, Catherine A.

; APPLICANT: Hickman, Robert K.

; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE

; FILE REFERENCE: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1

; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 58

; LENGTH: 215

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-8PL

US-08-911-824-58

## Query Match

Best Local Similarity 85.7%; Score 108; DB 8; Length 215;

Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTSVQWNST 22

||:|||||:|||||:|||||:  
Db 127 LLNLWGCKGRVLCYTSVKWNST 148

## RESULT 12

US-08-911-824-48

; Sequence 48, Application US/08911824

; Publication No. US20030004323A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Hackett, John R., Jr.

; APPLICANT: Yamaguchi, Julie

; APPLICANT: Golden, Alan M.

; APPLICANT: Brennan, Catherine A.

; APPLICANT: Hickman, Robert K.

; APPLICANT: Devare, Sushil G.

; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE

; FILE REFERENCE: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

; FILE REFERENCE: 6165.US.O1

; CURRENT APPLICATION NUMBER: US/08/911,824

; CURRENT FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 48

; LENGTH: 245

; TYPE: PRT

; ORGANISM: Human Immunodeficiency Virus

; FEATURE:

; OTHER INFORMATION: Encodes recombinant protein pGO-9PL

US-08-911-824-48

## Query Match

Best Local Similarity 85.7%; Score 108; DB 8; Length 245;

Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTSVQWNST 22

||:|||||:|||||:|||||:  
Db 127 LLNLWGCKGRVLCYTSVKWNST 148

## RESULT 13

US-08-911-824-120  
; Sequence 120, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 120  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL  
US-08-911-824-120

Query Match 85.7%; Score 108; DB 8; Length 281;  
Best Local Similarity 77.3%; Pred. No. 8.6e-07;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLSSWGCKGRVCYTSVQWNST 22  
||:|||||:|||||:  
Db 127 LLNLWGCKGRLCYTSVKWNET 148

## RESULT 14

US-08-911-824-52  
; Sequence 52, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 52  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL  
US-08-911-824-52

Query Match 85.7%; Score 108; DB 8; Length 373;  
Best Local Similarity 77.3%; Pred. No. 1.1e-06;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLSSWGCKGRVCYTSVQWNST 22  
||:|||||:|||||:  
Db 127 LLNLWGCKGRLCYTSVKWNET 148

## RESULT 15

US-08-911-824-60  
; Sequence 60, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 60  
; LENGTH: 460  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS  
US-08-911-824-60

Query Match 85.7%; Score 108; DB 8; Length 460;  
Best Local Similarity 77.3%; Pred. No. 1.4e-06;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLSSWGCKGRVCYTSVQWNST 22  
||:|||||:|||||:  
Db 372 LLNLWGCKGRLCYTSVKWNET 393

Search completed: May 7, 2004, 18:29:20  
Job time : 30.7815 secs

GREEN MOUNTAIN CORP.



Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQVNST 22  
||:||||||:||||:|

Db 607 LLNLWGCKRLLCYTSVKWNKT 628

RESULT 3

S22930

GP41 ENV protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C:Accession: S22930

R:Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de  
submitted to the EMBL Data Library, January 1995

A:Description: A novel HIV-1 O strain illustrates the diversity of the O group.

A:Reference number: S52929

A:Accession: S52930

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-104 <COH>

A:Cross-references: EMBL:X84328; NID:g95526; PIDN:CAA59066.1; PID:g95527

C:Superfamily: type E retrovirus env polyprotein

Query Match 80.2%; Score 101; DB 2; Length 104;  
Best Local Similarity 68.2%; Pred. No. 1.9e-07;  
Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQVNST 22  
||:||||||:||||:|

Db 45 LLNLWGCKRLLCYTSVQVNST 66

RESULT 4

S21990

env polyprotein gp120/gp41 - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 20

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000

C:Accession: S21990; S70423

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined

A:Reference number: S21990

A:Accession: S21990

A:Molecule type: DNA

A:Residues: 1-357 <STEL>

A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70423

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332, 'X', 334-357 <STE2>

A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176

C:Superfamily: type E retrovirus env polyprotein

Query Match 67.5%; Score 85; DB 2; Length 357;  
Best Local Similarity 59.1%; Pred. No. 0.00011;  
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQVNST 22  
||:||||||:||||:|

Db 93 LLGWCGRLLCYTSVQVNST 114

RESULT 5

S22002

env polyprotein gp120/gp41 - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 3L

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000

C:Accession: S22002; S70418

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined

A:Reference number: S21990

A:Accession: S22002

A:Molecule type: DNA

A:Residues: 1-358 <STEL>

A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70418

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333, 'X', 335-358 <STE2>

A:Cross-references: EMBL:X61352; NID:g60186

C:Superfamily: type E retrovirus env polyprotein

Query Match 67.5%; Score 85; DB 2; Length 358;  
Best Local Similarity 59.1%; Pred. No. 0.00011;  
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQVNST 22  
||:||||||:||||:|

Db 94 LLGWCGRLLCYTSVQVNST 115

RESULT 6

C41621

env polyprotein P - human immunodeficiency virus type 1 (fragment)

N:Alternate names: coat polyprotein

N:Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120

C:Species: human immunodeficiency virus type 1, HIV-1

A>Note: host Homo sapiens (man)

C>Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999

C:Accession: C41621

R:Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991

A>Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity

A:Reference number: A41621; MUID:92107924; PMID:1763038

A:Accession: C41621

A:Molecule type: DNA

A:Residues: 1-443 <BUR>

A:Cross-references: GB:M77230; NID:g328631; PIDN:AAB03792.1; PID:g555015

A>Note: this virus was isolated from the mother's sexual partner

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein

F:1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>

F:252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>

F:424-443/Domain: transmembrane #status predicted <TMN>

F:9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: ca

Query Match 67.5%; Score 85; DB 2; Length 443;  
Best Local Similarity 59.1%; Pred. No. 0.00013;  
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQVNST 22  
||:||||||:||||:|

Db 332 LLGWCGRLLCYTSVQVNST 353

RESULT 7

S54384

env polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Aug-1999

C:Accession: S54384

R:Theodore, T.; Buckler-White, A.J.



```
A;Accession: S22004
A:Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61353; NID:G60189; PIDN:CAA43618.1; PID:G60189
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid from patients with AIDS
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70419
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-292, X', 294-357 <STE2>
A;Cross-references: EMBL:X61353; NID:G60188
C;Superfamily: type E retrovirus env polyprotein

Query Match          65.1%; Score 82; DB 2; Length 357;
Best Local Similarity 54.5%; Pred.No. 0.00029;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LLSSWGCKGRLYCYTSVQVNST 22
        ||||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db       93 LGGIWGCSGKLICTTTPWNTS 114

RESULT 15
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S70422; S21996
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53
```



GenCore version S.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 5.54622 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-4  
Perfect score: 126  
Sequence: 1 LLSWGCKRLVCYTSVQWNST 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	67.5	853	1 ENV_HV122	P12487 human immun
2	85	67.5	854	1 ENV_SIVCZ	P17281 chimpanzee
3	85	67.5	855	1 ENV_HV126	P04580 human immun
4	82	65.1	843	1 ENV_HV1Y2	P35961 human immun
5	82	65.1	847	1 ENV_HV1S1	P19550 human immun
6	82	65.1	847	1 ENV_HV1W2	P05880 human immun
7	82	65.1	851	1 ENV_HV1B8	P04582 human immun
8	82	65.1	852	1 ENV_HV1B3	P12488 human immun
9	82	65.1	852	1 ENV_HV1S3	P19549 human immun
10	82	65.1	853	1 ENV_HV1MF	P19551 human immun
11	82	65.1	855	1 ENV_HV1A2	P03378 human immun
12	82	65.1	855	1 ENV_HV1OY	P20888 human immun
13	82	65.1	856	1 ENV_HV1B1	P03375 human immun
14	82	65.1	856	1 ENV_HV1H2	P04578 human immun
15	82	65.1	856	1 ENV_HV1LW	Q70626 human immun
16	82	65.1	856	1 ENV_HV1MN	P05877 human immun
17	82	65.1	856	1 ENV_HV1PV	P03376 human immun
18	82	65.1	856	1 ENV_HV1SC	P05878 human immun
19	82	65.1	856	1 ENV_HV1W1	P31872 human immun
20	82	65.1	856	1 ENV_HV1ZH	P05881 human immun
21	82	65.1	861	1 ENV_HV1BR	P03377 human immun
22	82	65.1	865	1 ENV_HV1RH	P04579 human immun
23	82	65.1	867	1 ENV_HV1J3	P12489 human immun
24	82	65.1	868	1 ENV_HV1C4	P05879 human immun
25	81	64.3	846	1 ENV_HV1ND	P18799 human immun
26	80	63.5	821	1 ENV_SIVGB	P22380 simian immun
27	80	63.5	848	1 ENV_HV1JR	P20871 human immun
28	80	63.5	856	1 ENV_HV1H3	P04624 human immun
29	78	61.9	853	1 ENV_HV1EL	P04581 human immun
30	78	61.9	863	1 ENV_HV1Z8	P05882 human immun
31	76	60.3	859	1 ENV_HV1MA	P04583 human immun
32	76	60.3	861	1 ENV_HV1KB	P31819 human immun
33	74	58.7	857	1 ENV_HV2KR	Q74126 human immun

34	72.5	57.5	856	1 ENV_HV2NZ	P05883 human immun
35	70	55.6	865	1 ENV_SIVAT	P05886 simian immun
36	69.5	55.2	712	1 ENV_HV2S2	P32536 human immun
37	69.5	55.2	846	1 ENV_HV2SB	P12449 human immun
38	69.5	55.2	859	1 ENV_HV2D2	P15831 human immun
39	69.5	55.2	859	1 ENV_HV2ST	P20872 human immun
40	69.5	55.2	885	1 ENV_SIVS4	P12492 simian immun
41	69	54.8	851	1 ENV_HV2D1	P17755 human immun
42	69	54.8	851	1 ENV_HV2G1	P18040 human immun
43	69	54.8	858	1 ENV_HV2RO	P04577 human immun
44	69	54.8	859	1 ENV_HV2CA	P24105 human immun
45	69	54.8	860	1 ENV_HV2BE	P18094 human immun

ALIGNMENTS

RESULT 1  
ENV\_HV122 STANDARD; PRT; 853 AA.  
AC P12487;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxId=11683;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Theodore T., Buckler-White A.;  
RL Submitted (NOV-1988) to the HIV data bank.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M22639; AAA45370.1; -  
DR PIR; S54384; S54384.  
DR HIV; M22639; ENV5226.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 31 BY SIMILARITY  
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 206 BY SIMILARITY.  
FT DISULFID 125 197 BY SIMILARITY.  
FT DISULFID 130 154 BY SIMILARITY.  
FT DISULFID 219 248 BY SIMILARITY.  
FT DISULFID 229 240 BY SIMILARITY.  
FT DISULFID 297 330 BY SIMILARITY.  
FT DISULFID 376 442 BY SIMILARITY.  
FT DISULFID 383 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).





```
Query Match          65.1%; Score 82; DB 1; Length 843;
Best Local Similarity 54.5%; Pred. No. 5.5e-05;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLYCVTSVQVNST 22
Db 579 LLGIWCSGKLICTTVPWNTS 600

RESULT 5
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65024; AAA45072.1; -.
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENVSSF162.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT CHAIN 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```
FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7F2F2AB CRC64;

Query Match          65.1%; Score 82; DB 1; Length 847;
Best Local Similarity 54.5%; Pred. No. 5.5e-05;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLYCVTSVQVNST 22
Db 583 LLGIWCSGKLICTTAVPWNAS 604

RESULT 6
ENV_HV1W2
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
at risk for AIDS.";
RL Science 232:1548-1553(1986).
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12507; AAB12990.1; -.
DR HIV; M12507; ENVSWMJ2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 202 BY SIMILARITY.
FT DISULFID 125 193 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 215 244 BY SIMILARITY.
FT DISULFID 225 236 BY SIMILARITY.
FT DISULFID 293 326 BY SIMILARITY.
```

```
FT DISULFID 372 435 BY SIMILARITY.
FT CARBOHYD 379 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96466 MW; CD1E33D73AA5BCAE CRC64;

Query Match 65.1%; Score 82; DB 1; Length 847;
Best Local Similarity 54.5%; Pred. No. 5.5e-05;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNST 22
DB 583 LLGIGCGSKGLICTTTPWPNAS 604

RESULT 7
ENV_HV1B8
ID ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
GN Human immunodeficiency virus type 1 (B8 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Pettway S.R., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; K02011; AAA44661.1; -.
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QO3; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENV$B8.
DR GlycoSuiteDB; P04582; -.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT CHAIN 1 30
FT CHAIN 31 506
FT DISULFID 54 74 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.
FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 65.1%; Score 82; DB 1; Length 851;
Best Local Similarity 54.5%; Pred. No. 5.6e-05;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTSVQWNST 22
DB 587 LLGIGCGSKGLICTTTPWPNAS 608

RESULT 8
ENV_HV1B8
ID ENV_HV1B8 STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
```



FT	CHAIN	31	509		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	510	853		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74		BY SIMILARITY.
FT	DISULFID	119	203		BY SIMILARITY.
FT	DISULFID	126	194		BY SIMILARITY.
FT	DISULFID	131	157		BY SIMILARITY.
FT	DISULFID	216	245		BY SIMILARITY.
FT	DISULFID	226	237		BY SIMILARITY.
FT	DISULFID	294	329		BY SIMILARITY.
FT	DISULFID	376	443		BY SIMILARITY.
FT	DISULFID	383	416		BY SIMILARITY.
FT	CARBOHYD	88	136		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	136	136		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	141	141		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	156	156		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	160	160		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186	186		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	195	195		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	232	232		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	260	260		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	274	274		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	287	287		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	293	293		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	299	299		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	330	330		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	334	354		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	384	384		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	390	390		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	395	395		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	404	404		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	446	446		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	461	461		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	609	609		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	614	614		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	623	623		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	635	635		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	672	672		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	748	748		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	814	814		N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	853 AA; 96912 MW; 3377899386F22ABA	CRC64;		
					Query Match 65.1%; Score 82; DB 1; Length 853;
					Best Local Similarity 54.5%; Pred. No. 5.6e-05;
					Matches 12; Conservative 5; Mismatches 5; Indels 0;
OY	1	LLSSWGCKGRLVCYTSTVQNNST 22			
Dd	590	LLGIWGCCKLICCTTAVPPNAS 611			
					:  :  :  :  :  :
					RESULT 11
ID	ENV_HVIIA2	STANDARD;	PRT;	855 AA.	
AC	P03378;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membran glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
DE	GNV.				
OS	Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).				
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				
OX	NCBI_Taxid=11685;				
RN	{1}				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85090453; PubMed=2578227;				
RA	Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,				
RA	Stempien M.M., Brown-Shmer S.L., Gee W.W., Renard A., Randolph				
RA	Levy J.A., Dina D., Luciw P.A.;				
RT	"Nucleotide sequence and expression of an AIDS-associated retrov				
RT	(ARV-2).";				
RI	Science. 227:484-492(1985)				









-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; K03455; AAB50262.1; -;  
 DR EMBL; AF038399; AAB99976.1; -;  
 DR EMBL; AF033819; AAC82596.1; -;  
 DR PDB; 1DF4; 26-JAN-00.  
 DR PDB; 1DF5; 26-JAN-00.  
 DR PDB; 1DLB; 02-APR-00.  
 DR PDB; 1G9M; 27-DEC-00.  
 DR PDB; 1GCI; 19-AUG-98.  
 DR PDB; 1GZL; 10-OCT-02.  
 DR PDB; 1K33; 10-OCT-01.  
 DR PDB; 1K34; 10-OCT-01.  
 DR HIV; K03455; ENVSHXB2.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 218 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 236 331 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;  
 65.1%; Score 82; DB 1; Length 856;  
 Query Match

Best Local Similarity 54.5%; Pred. No. 5.6e-05;  
 Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 LLSWGCKGRVLCYTSVQWNST 22  
 DB 592 LLIWGCGSKLICITTAVPWNAS 613  
 RESULT 15  
 ENV\_HVILW STANDARD; PRT; 856 AA.  
 AC Q70626;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (LMI2.3 isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=82834;  
 RN [1]\_TaxID=82834;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95127297; PubMed=7826699;  
 RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
 Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
 "Viral variability and serum antibody response in a laboratory worker  
 RT infected with HIV type 1 (HIV type IIB).";  
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U12055; AAA76690.1; -;  
 DR PDB; 1IF3; 02-MAY-01.  
 DR GlycoSuiteDB; Q70626; -;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 218 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;  
 65.1%; Score 82; DB 1; Length 856;  
 Query Match

FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA;	96938 MW;	0C241332CF7E6687	CRC64;

Query Match 65.1%; Score 82; DB 1; Length 856;

Best Local Similarity 54.5%; Pred. No. 5.6e-05;

Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGELVCYTSVQWNST 22

Db 592 LLGIWGCKGLICTTAVPMNAS 613

Search completed: May 7, 2004, 17:43:55

Job time : 5.54622 secs

COPIES OF THE  
RECORDS OF THE  
FEDERAL BUREAU OF  
INVESTIGATION  
AND  
THE  
DEPARTMENT OF JUSTICE  
ARE AVAILABLE FOR  
REVIEW AND  
REPRODUCTION  
BY THE PUBLIC  
AT THE  
FEDERAL BUREAU OF  
INVESTIGATION  
AND  
THE  
DEPARTMENT OF JUSTICE  
LIBRARY  
1000 PENNSYLVANIA AVENUE, N.W.  
WASHINGTON, D.C. 20535  
(202) 452-5000

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 27.916 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-4

Perfect score: 126

Sequence: 1 LLSSWGCKGRLVCYTSVQWNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTEMBL 25:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phase:\*
- 11: sp\_plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp\_vertebrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_rvirus:\*
- 17: sp\_bacteriaph:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	92.1	219	15 Q9IEB6	Q9IEB6 human immun
2	115	91.3	116	15 Q7ZJN9	Q7ZJN9 human immun
3	115	91.3	135	15 Q9DQL9	Q9DQL9 human immun
4	115	91.3	216	15 Q9IEC5	Q9IEC5 human immun
5	115	91.3	218	15 Q9IEE9	Q9IEE9 human immun
6	115	91.3	242	15 Q9IE31	Q9IE31 human immun
7	115	91.3	538	15 Q9IED5	Q9IED5 human immun
8	113	89.7	214	15 Q9DIK3	Q9DIK3 human immun
9	113	89.7	216	15 Q9IEC7	Q9IEC7 human immun
10	113	89.7	234	15 Q9IEA6	Q9IEA6 human immun
11	113	89.7	535	15 Q9IEF2	Q9IEF2 human immun
12	113	89.7	551	15 Q9IEE1	Q9IEE1 human immun
13	112	88.9	116	15 Q11941	Q11941 human immun
14	112	88.9	116	15 Q0458	Q0458 human immun
15	112	88.9	130	15 Q9IHU9	Q9IHU9 human immun
16	112	88.9	158	15 Q8J3N6	Q8J3N6 human immun

17	112	88.9	183	15 Q9IEC1	Q9IEC1 human immun
18	112	88.9	220	15 Q9IEC9	Q9IEC9 human immun
19	112	88.9	235	15 Q9IE54	Q9IE54 human immun
20	112	88.9	512	15 Q9IED2	Q9IED2 human immun
21	112	88.9	872	15 Q8Q7H0	Q8Q7H0 human immun
22	112	88.9	879	15 Q9WIU9	Q9WIU9 human immun
23	112	88.9	880	15 Q8Q7I9	Q8Q7I9 human immun
24	112	88.9	882	15 Q8Q7F9	Q8Q7F9 human immun
25	112	88.9	887	15 Q8Q7H6	Q8Q7H6 human immun
26	112	88.9	887	15 Q8Q7G9	Q8Q7G9 human immun
27	112	88.9	900	15 Q9QN28	Q9QN28 human immun
28	111	88.1	114	15 Q0448	Q0448 human immun
29	111	88.1	209	15 Q9IE66	Q9IE66 human immun
30	110	87.3	118	15 Q0451	Q0451 human immun
31	110	87.3	214	15 Q9IE96	Q9IE96 human immun
32	110	87.3	216	15 Q9IEA5	Q9IEA5 human immun
33	110	87.3	219	15 Q9IEC8	Q9IEC8 human immun
34	110	87.3	242	15 Q9IE30	Q9IE30 human immun
35	110	87.3	544	15 Q9IED9	Q9IED9 human immun
36	110	87.3	548	15 Q9IED6	Q9IED6 human immun
37	110	87.3	865	15 Q8Q7H7	Q8Q7H7 human immun
38	110	87.3	867	15 Q8Q7G8	Q8Q7G8 human immun
39	109	86.5	116	15 Q0459	Q0459 human immun
40	109	86.5	118	15 Q9E5S0	Q9E5S0 human immun
41	109	86.5	119	15 Q9IHU5	Q9IHU5 human immun
42	109	86.5	121	15 Q9IHV7	Q9IHV7 human immun
43	109	86.5	134	15 Q9IHV4	Q9IHV4 human immun
44	109	86.5	137	15 Q9IHV5	Q9IHV5 human immun
45	109	86.5	146	15 Q9WRV2	Q9WRV2 human immun

ALIGNMENTS

RESULT 1

Q9IEB6 PRELIMINARY; PRT; 219 AA.

AC Q9IEB6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF14;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236404; CAB96252.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 92.1%; Score 116; DB 15; Length 219;

Best Local Similarity 86.4%; Pred. No. 2.3e-10;

Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22

||:|||||:|||||:|

Db 57 LLSSWGCKGRLVCYTSVKWNNT 78

```
RESULT 2
Q7ZJN9
ID Q7ZJN9 PRELIMINARY; PRT; 116 AA.
AC Q7ZJN9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08692A;
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol
RT integrase, and env gp41.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY214120; AAO61840.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match 91.3%; Score 115; DB 15; Length 116;
Best Local Similarity 86.4%; Pred. No. 1.7e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
DB 32 LLSSWGCKGRLVCYTSVRKNQT 53

RESULT 3
Q9DQL9
ID Q9DQL9 PRELIMINARY; PRT; 135 AA.
AC Q9DQL9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20584646; PubMed=11153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,
RA Vadiello J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).
DR EMBL; AF255939; AAG36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 135 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match 91.3%; Score 115; DB 15; Length 135;
Best Local Similarity 81.8%; Pred. No. 2e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
```

```
Db 48 LLNSWGCKGRLVCYTSVKWNST 69

RESULT 4
Q9IEC5
ID Q9IEC5 PRELIMINARY; PRT; 216 AA.
AC Q9IEC5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF07;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236394; CAB96243.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 91.3%; Score 115; DB 15; Length 216;
Best Local Similarity 86.4%; Pred. No. 3.2e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
DB 47 LLNSWGCKGRLVCYTSVKWNST 68

RESULT 5
Q9IE95
ID Q9IE95 PRELIMINARY; PRT; 218 AA.
AC Q9IE95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF51;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236425; CAB96273.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 218 218
SQ SEQUENCE 218 AA; 25243 MW; E7F0E1E20E79FA8 CRC64;
```

```

DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
DR FT NON_TER 1 538
DR FT NON_TER 538
DR SQ SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEF1D CRC64;

Query Match 91.3%; Score 115; DB 15; Length 538;
Best Local Similarity 81.8%; Pred. No. 8.1e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWNST 22
|||||:|:|:|:|:|:|:|
Db 374 LLSSWGCKGRLVCYTSVKWNTT 395

RESULT 8
Q9DIK3 PRELIMINARY; PRT; 214 AA.
ID AC Q9DIK3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 protein (Fragment).
DE ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF120;
RA Ropes P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA Depienne C., Brun-Verinet F., Dormont D., Simon F.;
RT "Phylogenetic analysis of 49 newly derived HIV-1 group O strains: High
RL viral diversity but no group M-like subtypes.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ298125; CAC18806.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR KW Transmembrane.
DR FT NON_TER 1 214
DR FT NON_TER 214 214
DR SQ SEQUENCE 214 AA; 24436 MW; FFAB9D08F861ABB CRC64;

Query Match 89.7%; Score 113; DB 15; Length 214;
Best Local Similarity 81.8%; Pred. No. 6.6e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRLVCYTSVQWNST 22
|||||:|:|:|:|:|:|:|
Db 45 LLNSWGCKGRLVCYTSVKWNTT 66

RESULT 9
Q9IEC7 PRELIMINARY; PRT; 216 AA.
ID AC Q9IEC7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
DE ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF03;
RA Ropes P., Robertson D., Sandrine S., Christel D., Francoise S.

```

```
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236392; CAB96241.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25015 MW; 780C5F39CA6997C2 CRC64;

Query Match 89.7%; Score 113; DB 15; Length 216;
Best Local Similarity 86.4%; Pred. NO. 6.6e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 55 LLSWGCKGRLVCYTSVKWNT 76

RESULT 10
Q9IEA6 PRELIMINARY; PRT; 234 AA.
AC Q9IEA6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=YBF22;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236414; CAB96262.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 234
SQ SEQUENCE 234 AA; 26737 MW; 34CFDBE999DEFB82 CRC64;

Query Match 89.7%; Score 113; DB 15; Length 234;
Best Local Similarity 86.4%; Pred. NO. 7.2e-10;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 62 LLSWGCKGRLVCYTSVKWNT 83

RESULT 11
Q9IEF2 PRELIMINARY; PRT; 535 AA.
AC Q9IEF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
```

```
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF03;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
Deplienne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133055; CAB96216.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 535
SQ SEQUENCE 535 AA; 59682 MW; 7234BE8378DD12C5 CRC64;

Query Match 89.7%; Score 113; DB 15; Length 535;
Best Local Similarity 86.4%; Pred. NO. 1.7e-09;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 374 LLSWGCKGRLVCYTSVKWNT 395

RESULT 12
Q9IEE1 PRELIMINARY; PRT; 551 AA.
AC Q9IEE1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF22;
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
Deplienne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133066; CAB96227.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 551
SQ SEQUENCE 551 AA; 61704 MW; 1CF824CAD5A89ECF CRC64;

Query Match 89.7%; Score 113; DB 15; Length 551;
Best Local Similarity 86.4%; Pred. NO. 1.7e-09;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 379 LLSWGCKGRLVCYTSVKWNT 400
```



```
RESULT 13
ID O11941 PRELIMINARY; PRT; 116 AA.
AC O11941;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ABT124;
RX MEDLINE=97340911; PubMed=9197385;
RA Brennan C.A., Hackett J. Jr., Zekeng L., Lund J.K., Vallari A.S.,
RA Hickman R.K., Gurtler L., Kaptue L., von Overbeck J., Hampl H.,
RA Devare S.G.;
RT "Sequence of gp41env immunodominant region of HIV type 1 group O from
RT west central Africa.";
RL AIDS Res. Hum. Retroviruses 13:901-904(1997).
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 13789 MW; 488A9A40F4255E3E CRC64;

Query Match 88.9%; Score 112; DB 15; Length 116;
Best Local Similarity 81.8%; Pred. No. 5.1e-10;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
||:|||||:|||||:|||||
Db 39 LLNLWGCKGRLVCYTSVQWNST 60

RESULT 14
ID O40458 PRELIMINARY; PRT; 116 AA.
AC O40458;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope transmembrane glycoprotein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=group O;
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;
RT "Molecular characterization of envelope transmembrane glycoprotein of
RT 14 new human immunodeficiency virus type 1 group O strains from
RT different African countries.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09774; CAA70913.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 13865 MW; 540F5B3B7AD849D8 CRC64;

Query Match 88.9%; Score 112; DB 15; Length 116;
```

```
Best Local Similarity 81.8%; Pred. No. 5.1e-10;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
||:|||||:|||||:|||||
Db 39 LLNLWGCKGRLVCYTSVQWNKT 60

RESULT 15
ID Q91HU9 PRELIMINARY; PRT; 130 AA.
AC Q91HU9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Fieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF29235; AAF71912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match 88.9%; Score 112; DB 15; Length 130;
Best Local Similarity 81.8%; Pred. No. 5.7e-10;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTSVQWNST 22
||:|||||:|||||:|||||
Db 40 LLNLWGCKGRLVCYTSVQWNKT 61

Search completed: May 7, 2004, 17:49:12
Job time : 28.916 secs
```

SECRET  
NOV 16 1954  
U.S. DEPT. OF STATE

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 41.8739 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-5

Perfect score: 127

Sequence: 1 LIQSWGCKGRLVCYTSVQMNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	100.0	22	2 AAW80463	AAW80463 Peptide d
2	122	96.1	22	2 AAW80462	AAW80462 Peptide d
3	118	92.9	22	2 AAW80461	AAW80461 Peptide d
4	115	90.6	22	2 AAW80466	AAW80466 Peptide d
5	114	89.8	40	2 AAW07346	AAW07346 Partial s
6	112	88.2	40	2 AAW07352	AAW07352 Partial s
7	111	87.4	32	2 AAW80469	AAW80469 Peptide d
8	111	87.4	41	2 AAW07351	AAW07351 Partial s
9	111	87.4	113	2 AAY05565	AAY05565 HIV-1 gro
10	111	87.4	715	2 AAY05625	AAY05625 HIV-1 gro
11	110	86.6	33	3 AAB12212	AAB12212 Partial s
12	110	86.6	113	2 AAY05559	AAY05559 HIV-1 gro
13	109	85.8	22	2 AAW80460	AAW80460 Peptide d
14	109	85.8	33	3 AAB12231	AAB12231 Partial s
15	109	85.8	40	2 AAW07344	AAW07344 Partial s
16	109	85.8	40	2 AAW07343	AAW07343 Partial s
17	109	85.8	41	2 AAW07353	AAW07353 Partial s
18	109	85.8	117	2 AAY05548	AAY05548 HIV-1 gro
19	108	85.0	116	2 AAY05555	AAY05555 HIV-1 gro
20	107	84.3	22	2 AAW80464	AAW80464 Peptide d
21	107	84.3	23	3 AAY05623	AAY05623 HIV-1 gro
22	107	84.3	33	3 AAB12236	AAB12236 Partial s
23	107	84.3	113	2 AAY05551	AAY05551 HIV-1 gro
24	107	84.3	113	2 AAY05550	AAY05550 HIV-1 gro
25	107	84.3	200	3 AAY77373	AAW77373 HIV-1 gro

26	107	84.3	215	2 AAY09499	Aay09499 HIV-1 Gro
27	107	84.3	215	2 AAY06983	Aay06983 Recombina
28	107	84.3	215	3 AAY77374	Aay77374 HIV-1 gro
29	107	84.3	245	2 AAY09493	Aay09493 HIV-1 Gro
30	107	84.3	245	2 AAY06977	Aay06977 Recombina
31	107	84.3	245	3 AAY77369	Aay77369 HIV-1 gro
32	107	84.3	281	2 AAY09507	Aay09507 HIV-1 Gro
33	107	84.3	373	2 AAY09495	Aay09495 HIV-1 Gro
34	107	84.3	373	2 AAY06979	Aay06979 Recombina
35	107	84.3	460	2 AAY09500	Aay09500 HIV-1 Gro
36	107	84.3	460	2 AAY06984	Aay06984 Recombina
37	107	84.3	460	3 AAY77375	Aay77375 HIV-1 gro
38	107	84.3	474	3 AAY77371	Aay77371 HIV-1 gro
39	107	84.3	488	2 AAY09504	Aay09504 HIV-1 Gro
40	107	84.3	490	2 AAY09494	Aay09494 HIV-1 Gro
41	107	84.3	490	2 AAY06978	Aay06978 Recombina
42	107	84.3	490	3 AAY77370	Aay77370 HIV-1 gro
43	107	84.3	526	2 AAY09505	Aay09505 HIV-1 Gro
44	107	84.3	618	2 AAY09496	Aay09496 HIV-1 Gro
45	107	84.3	618	2 AAY06980	Aay06980 Recombina

ALIGNMENTS

RESULT 1

AAW80463

ID AAW80463 standard; peptide; 22 AA.

XX AC AAW80463;

XX DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)

XX DE Peptide derived from a conserved sequence of group O human HIV.

XX KW Group O human immune deficiency virus; HIV; detection; infection.

XX OS Synthetic.

OS Human immunodeficiency virus 1.

XX PN WO9845323-A1.

XX PD 15-OCT-1998.

XX PF 06-APR-1998; 98WO-FR000691.

XX PR 09-APR-1997; 97FR-00004356.

XX PR 24-FEB-1998; 98FR-00002212.

XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX PI Cheneboux DMB, Delagneau JH, Gadelle SXJ, Rieunier FY;

XX XX WPI; 1998-583190/49.

XX DR New synthetic peptide(s) - useful for, e.g. detecting infection by human immune deficiency virus of group O.

XX PT Claim 6; Page 42; 55pp; French.

XX PS AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences connected around short highly conserved sequences present in isolates of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)

XX CC (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 22 AA;

Query Match 100.0%; Score 127; DB 2; Length 22;



PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 PS  
 XX Claim 6; Page 43; 55pp; French.  
 XX  
 CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 SQ Sequence 22 AA;  
 Query Match 90.6%; Score 115; DB 2; Length 22;  
 Best Local Similarity 86.4%; Pred. No. 2.9e-08;  
 Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSWGCKGRLVCYTSVQWNST 22  
 DB 1 LLSWGCKGRLVCYTSVQWNST 22  
 RESULT 5  
 AAW07346  
 ID AAW07346 standard; peptide; 40 AA.  
 XX  
 AC AAW07346;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95FR-00002236.  
 XX  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;  
 DR WPI; 1996-412779/41.  
 XX  
 DR N-PSDB; AAT44922.  
 XX  
 PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 34; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene

(env): group M containing sub-groups A-G, and group O containing the strains ANT70 and MVP5180. The invention relates to the discovery of several new strains of HIV-1 which can be placed in group O, based on the partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39 and AAW07329-64). The novel strains have been deposited as retroviruses CCMC I-1544 (BCF02 (SSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547 (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein encoded by the env gene. The nucleic acids can be used to detect gp. O HIV-1 strains by hybridisation or (as primers) by gene amplification. CC also for screening and typing of such strains. Peptides encoded by the nucleic acids can be used as immunogens to raise Ab for detecting gp. O HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 40 AA;  
 Query Match 89.8%; Score 114; DB 2; Length 40;  
 Best Local Similarity 86.4%; Pred. No. 6.9e-08;  
 Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSWGCKGRLVCYTSVQWNST 22  
 DB 19 LLSWGCKGRLVCYTSVQWNST 40  
 RESULT 6  
 AAW07352  
 ID AAW07352 standard; peptide; 40 AA.  
 XX  
 AC AAW07352;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF13.  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95FR-00002236.  
 XX  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Lousertajaka I, Ly T, Chaixbaudier M;  
 DR WPI; 1996-412779/41.  
 XX  
 DR New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 46; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene (env): group M containing sub-groups A-G, and group O containing the strains ANT70 and MVP5180. The invention relates to the discovery of several new strains of HIV-1 which can be placed in group O, based on the partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39 and AAW07329-64). The novel strains have been deposited as retroviruses



AA05565 standard; protein; 113 AA.

AA05565;

17-OCT-2003 (revised)

19-JUL-1999 (first entry)

HIV-1 group O isolate MP539-PBMC gp41 antigen.

HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine; diagnosis; AIDS.

Human immunodeficiency virus 1.

Key Location/Qualifiers

Misc-difference 65 /note= "encoded by AMW"

Misc-difference 74 /note= "encoded by ATR"

Misc-difference 84 /note= "encoded by GAK"

Misc-difference 86 /note= "encoded by AGY"

WO9904011-A2.

28-JAN-1999.

20-JUL-1998; 98WO-EP004522.

18-JUL-1997; 97EP-00870110.

(INNO-) INNOGENETICS NV.

Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;

WPI; 1999-132255/11.

N-PSDB; AA05565.

New isolated HIV-1 group O strains - used to produce polynucleotides, antigens and antibodies for use in diagnosis and in vaccines for prevention of HIV-1 infection.

Claim 3; Fig 6; 162pp; English.

The present sequence is an antigen of the gp41 protein of HIV-1 group O (Outlier) strain MP539-PBMC, a Cameroon isolate. The invention relates to new HIV-1 group O antigens (see AA05546-625), and the use of these antigens, or nucleic acids encoding them (see AA05546-625), in the diagnosis and prophylaxis of AIDS. They can be used as reagents for detecting HIV-1 group O infection and for differentiating different types of HIV-1 group O infection. Vaccines that provide protective immunity against HIV-1 infection, in particular against HIV-1 group O infection, comprise at least one HIV-1 type O antigen, a nucleic acid encoding such an antigen, a virus-like particle comprising such an antigen, or an attenuated form of an HIV-1 type O strain. The invention also relates to new HIV-1 group O strains, obtained from patients from Cameroon, Gabon, Tchad, Nigeria, Senegal and Niger. (Updated on 17-OCT-2003 to standardise OS field)

Query Match 87.4%; Score 111; DB 2; Length 113;

Best Local Similarity 81.8%; Pred. No. 4.5e-07;

Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRLVCYTSVQWNST 22

Db 39 LLNLWGCKGRLVCYTSVQWNKT 60

RESULT 10

AA05565

Partial sequence of HIV-1 strain BCF13 gp41 immunodominant region.

KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
KW acquired immunodeficiency syndrome; group O HIV; gp41; BCF13.  
OS Human immunodeficiency virus 1.  
PN EP1013766-A2.  
XX  
PD 28-JUN-2000.  
XX  
XX 29-NOV-1999; 99EP-00309491.  
XX  
XX 30-NOV-1998; 98US-0110292P.  
PR 08-FEB-1999; 99US-0119138P.  
PR 04-NOV-1999; 99US-00433428.  
XX  
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
XX  
XX De Leys R, Zheng J;  
XX WPI; 2000-402205/35.  
DR  
XX New antigenic peptides and peptide functional derivatives, useful for  
PT detection of antibodies produced in response to human immunodeficiency  
PT virus group O antibodies.  
XX  
PS Example 1; Fig 1; 52pp; English.  
XX  
CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
CC Virus type 1 (HIV-1) strain BCF13. HIV is the principle aetiological  
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
CC envelope protein, and may be used as an antigen for the detection of  
CC antibodies produced in response to HIV infection. BCF13 is a member of  
CC HIV group O (outlier). The present sequence is the immunodominant region  
CC of gp41. This sequence was used in a sequence homology alignment, which  
CC in turn was used to derive a consensus sequence peptide: peptide 147  
CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 33 AA;  
Query Match 86.6%; Score 110; DB 3; Length 33;  
Best Local Similarity 90.0%; Pred. No. 1.9e-07;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LQSWGCKGRVCYTSVQWN 20  
DB 14 LLNSWGCKGRVCYTSVEWN 33  
RESULT 12  
AAV05559  
ID AAY05559 standard; protein; 113 AA.  
XX  
XX AAY05559;  
AC  
XX 17-OCT-2003 (revised)  
DT 19-JUL-1999 (first entry)  
XX  
XX HIV-1 group O isolate BSD189 gp41 antigen.  
DE  
XX HIV-1 group O; Outlier strain; gp41; envelope protein; antigen; vaccine;  
KW diagnosis; AIDS.  
KW  
XX Human immunodeficiency virus 1.  
OS  
XX WO9904011-A2.  
PN  
XX 28-JAN-1999.  
PD  
XX 20-JUL-1998; 98WO-EP004522.  
PF  
XX 18-JUL-1997; 97EP-00870110.  
PR  
XX (INNO-) INNOGENETICS NV.  
PA

XX Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;  
PI WPI; 1999-132255/11.  
DR N-PSDB; AAX25167.  
XX  
XX New isolated HIV-1 group O strains - used to produce polynucleotides,  
PT antigens and antibodies for use in diagnosis and in vaccines for  
PT prevention of HIV-1 infection.  
PT  
XX Claim 3; Fig 6; 162pp; English.  
PS  
XX The present sequence is an antigen of the gp41 protein of HIV-1 group O  
CC (Outlier) strain BSD189, a Cameroon isolate. The invention relates to new  
CC HIV-1 group O antigens (see AAY05546-625), and the use of these antigens,  
CC or nucleic acids encoding them (see AAX25154-80), in the diagnosis and  
CC prophylaxis of AIDS. They can be used as reagents for detecting HIV-1  
CC group O infection and for differentiating different types of HIV-1 group  
CC O infection. Vaccines that provide protective immunity against HIV-1  
CC infection, in particular against HIV-1 group O infection, comprise at  
CC least one HIV-1 type O antigen, a nucleic acid encoding such an antigen,  
CC a virus-like particle comprising such an antigen, or an attenuated form  
CC of an HIV-1 type O strain. The invention also relates to new HIV-1 group  
CC O strains, mostly from patients from Cameroon and its neighbouring  
CC countries. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 113 AA;  
Query Match 86.6%; Score 110; DB 2; Length 113;  
Best Local Similarity 81.8%; Pred. No. 6.1e-07;  
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LQSWGCKGRVCYTSVQWNST 22  
DB 39 LLNLWGCKGRVCYTSVQWNMT 60  
RESULT 13  
AAW80460  
ID AAW80460 standard; peptide; 22 AA.  
XX  
XX AAW80460;  
AC  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
XX Peptide derived from a conserved sequence of group O human HIV.  
DE  
XX Group O human immune deficiency virus; HIV; detection; infection.  
KW  
XX Synthetic.  
XX Human immunodeficiency virus 1.  
OS  
XX WO9845323-A1.  
PN  
XX 15-OCT-1998.  
PD  
XX 06-APR-1998; 98WO-FR000691.  
PF  
XX 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
PR  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
PA  
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
XX WPI; 1998-583190/49.  
DR  
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
PT  
XX Claim 6; Page 42; 55pp; French.  
PS



XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 22 AA;  
 SQ

Query Match 85.8%; Score 109; DB 2; Length 22;  
 Best Local Similarity 81.8%; Pred. No. 1.8e-07;  
 Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLQSWGCKGRLVCYTSVQWNST 22  
 |||:|||||:  
 Db 1 LLSLWGCKGRLVCYTSVQWNST 22

RESULT 14  
 AAB12231  
 ID AAB12231 standard; peptide; 33 AA.  
 XX AAB12231;  
 AC  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 10-NOV-2000 (first entry)  
 XX  
 XX  
 DE Partial sequence of HIV-1 strain MAN gp41 immunodominant region.  
 XX  
 KW HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; MAN.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN EP1013766-A2.  
 XX  
 PD 28-JUN-2000.  
 XX  
 PF 29-NOV-1999; 99BP-00309491.  
 XX  
 PR 30-NOV-1998; 98US-0110292P.  
 PR 08-FEB-1999; 99US-0119138P.  
 PR 04-NOV-1999; 99US-00433428.  
 XX  
 XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 PA  
 PI De Lays R, Zheng J;  
 XX  
 XX WPI; 2000-402205/35.  
 DR  
 XX  
 PT New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.  
 XX  
 XX Example 1; Fig 1; 52pp; English.  
 PS  
 XX  
 CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain MAN. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. MAN is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 XX Sequence 33 AA;  
 SQ

Query Match 85.8%; Score 109; DB 3; Length 33;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-07;  
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLQSWGCKGRLVCYTSVQWN 20  
 |||:|||||:  
 Db 14 LLSLWGCKGRLVCYTSVKWN 33

RESULT 15  
 AAW07344  
 ID AAW07344 standard; peptide; 40 AA.  
 XX  
 XX AAW07344;  
 AC  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF01 (FAN).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 XX 26-FEB-1996; 96WO-FR000294.  
 PF  
 XX 27-FEB-1995; 95FR-00002236.  
 PR  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 PA  
 XX Simon F, Saragosti S, Loussertajaka I, Ly T, Chaixbaudier M;  
 PI  
 XX WPI; 1996-412779/41.  
 DR  
 DR N-PSDB; AAT44923.  
 XX  
 XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 XX Claim 12; Page 33; 71pp; French.  
 PS  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MVP5180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-84). The novel strains have been deposited as retroviruses  
 CC CINC I-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF01 (FAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification.  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 XX Sequence 40 AA;  
 SQ

Query Match 85.8%; Score 109; DB 2; Length 40;  
 Best Local Similarity 81.8%; Pred. No. 3.1e-07;  
 Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLQSWGCKGRLVCYTSVQWNST 22  
 |||:|||||:  
 Db 19 LLSLWGCKGRLVCYTSVKWNST 40

Search completed: May 7, 2004, 17:42:43  
Job time : 41.8739 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 12.1092 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-5

Perfect score: 127

Sequence: 1 LLOSMCKGRLVCYTSVQMNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCITUS COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	89.8	40	3	US-08-894-699-39
2	114	89.8	40	3	US-09-444-410-39
3	112	88.2	40	3	US-08-894-699-68
4	112	88.2	40	3	US-09-444-410-68
5	111	87.4	41	3	US-08-894-699-67
6	111	87.4	41	3	US-09-444-410-67
7	111	87.4	41	3	US-08-894-699-37
8	111	87.4	41	3	US-09-444-410-36
9	110	86.6	33	3	US-09-433-428D-6
10	110	86.6	113	4	US-09-462-917A-28
11	109	85.8	33	3	US-09-433-428D-25
12	109	85.8	40	3	US-08-894-699-36
13	109	85.8	40	3	US-08-894-699-37
14	109	85.8	40	3	US-09-444-410-37
15	109	85.8	41	3	US-08-894-699-69
16	109	85.8	41	3	US-09-444-410-69
17	109	85.8	41	3	US-08-894-699-69
18	109	85.8	117	4	US-09-462-917A-6
19	108	85.0	116	4	US-08-894-699-39
20	107	84.3	23	4	US-09-462-917A-137
21	107	84.3	33	3	US-09-433-428D-30
22	107	84.3	113	4	US-09-462-917A-10
23	107	84.3	113	4	US-09-462-917A-12
24	107	84.3	215	2	US-08-912-129A-58
25	107	84.3	245	2	US-08-912-129A-48
26	107	84.3	373	2	US-08-912-129A-52
27	107	84.3	460	2	US-08-912-129A-60

28 107 84.3 490 2 US-08-912-129A-50 Sequence 50, Appl  
29 107 84.3 618 2 US-08-912-129A-54 Sequence 54, Appl  
30 107 84.3 873 2 US-08-912-129A-61 Sequence 61, Appl  
31 106 83.5 23 4 US-09-462-917A-95 Sequence 95, Appl  
32 106 83.5 33 3 US-09-433-428D-8 Sequence 8, Appl  
33 106 83.5 40 3 US-08-894-699-40 Sequence 40, Appl  
34 106 83.5 40 3 US-09-444-410-40 Sequence 40, Appl  
35 106 83.5 110 4 US-09-462-917A-14 Sequence 14, Appl  
36 106 83.5 113 4 US-09-462-917A-2 Sequence 2, Appl  
37 106 83.5 113 4 US-09-462-917A-18 Sequence 18, Appl  
38 106 83.5 115 4 US-09-462-917A-24 Sequence 24, Appl  
39 105 82.7 110 4 US-09-462-917A-16 Sequence 16, Appl  
40 104 81.9 24 4 US-09-462-917A-138 Sequence 138, Appl  
41 104 81.9 37 4 US-08-817-441-86 Sequence 86, Appl  
42 104 81.9 40 3 US-08-894-699-41 Sequence 41, Appl  
43 104 81.9 40 3 US-08-894-699-42 Sequence 42, Appl  
44 104 81.9 40 3 US-09-444-410-41 Sequence 41, Appl  
45 104 81.9 40 3 US-09-444-410-42 Sequence 42, Appl

## ALIGNMENTS

RESULT 1  
US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear



```

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLSCULE TYPE: peptide
; US-08-894-699-67

Query Match 87.4%; Score 111; DB 3; Length 41;
Best Local Similarity 81.8%; Pred. No. 2e-09;
Matches 18; Conservative 2; Mismatches 2; Indels

QY 1 LQSWGCKGRLCVYTSVQWNST 22
   || |||||:||||:||||
Db 19 LNLWGCKGRLCVTSVKWNST 40

RESULT 6
US-09-444-410-67
; Sequence 67, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIK-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/894,699  
; FILING DATE:  
; PRIOR APPLICATION DATA: FR 95/02236  
; APPLICATION NUMBER: 27-FEB-1995  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 41 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-444-410-67

Query Match 87.4%; Score 111; DB 3; Length 41;  
Best Local Similarity 81.8%; Pred. No. 2e-09;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLOSGCKGRVCYTSVQWNST 22  
Db 19 LLNLWGCKGRLCYTSVKWNST 40

## RESULT 7

US-09-462-917A-40  
; Sequence 40, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc.feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-09-462-917A-40

Query Match 87.4%; Score 111; DB 4; Length 113;  
Best Local Similarity 81.8%; Pred. No. 5.7e-09;  
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLOSGCKGRVCYTSVQWNST 22  
Db 39 LLNLWGCKGRLCYTSVQWNT 60

## RESULT 8

US-09-462-917A-134  
; Sequence 134, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine

; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 134  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Human  
US-09-462-917A-134

Query Match 87.4%; Score 111; DB 4; Length 715;  
Best Local Similarity 86.4%; Pred. No. 3.8e-08;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLOSGCKGRVCYTSVQWNST 22  
Db 628 LLNLWGCKGRVCYTSVQWNET 649

## RESULT 9

US-09-433-428D-6  
; Sequence 6, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-6

Query Match 86.6%; Score 110; DB 3; Length 33;  
Best Local Similarity 90.0%; Pred. No. 2.2e-09;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLOSGCKGRVCYTSVQW 20  
Db 14 LLNLWGCKGRVCYTSVEWN 33

## RESULT 10

US-09-462-917A-28  
; Sequence 28, Application US/09462917A  
; Patent No. 6511801  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014 11362.0014.NPUS00  
; CURRENT APPLICATION NUMBER: US/09/462,917A  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 113  
; TYPE: PRT

```

/ FILING DATE: 26-FEB-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 95/02236
/ FILING DATE: 27-FEB-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: OBLON, NORMAN F.
/ REGISTRATION NUMBER: 24,614
/ REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-413-3000
/ TELEFAX: 703-413-2220
/ INFORMATION FOR SEQ ID NO: 36:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/
; US-08-894-699-36
;
; Query Match 85.8%; Score 109; DB 3; Length 40;
; Best Local Similarity 81.8%; Pred. No. 3.8e-09;
; Matches 18; Conservative 2; Mismatches 2; Indels
;
QY 1 LLQSWGCKGRIVCYTSVQNNST 22
; ||| |||||:|||||:|
;
Db 19 LLNSWGCKGRIVCYTSVKWNWT 40
;
;
; RESULT 13
; US-08-894-699-37
; Sequence 37, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSEST-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: FLOOR
; STATE: ARLINGTON
; COUNTRY: VA
; COUNTRY: USA
; ZIP: 22202
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220

```

; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-37

Query Match 85.8%; Score 109; DB 3; Length 40;  
Best Local Similarity 81.8%; Pred. No. 3.8e-09;  
Matches 18; Conservative 2; Mismatches 2; Indels 0;

Qy 1 LLQSWGCKGRLVCYTSVQWNST 22  
Db 19 LLNLWGCKGRLLCYTSVKWNST 40

RESULT 14

US-09-444-410-36  
; Sequence 36, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/444,410  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/894,699  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-444-410-36

Query Match 85.8%; Score 109; DB 3; Length 40;  
Best Local Similarity 81.8%; Pred. No. 3.8e-09;

Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 LLQSWGCKGRLVCYTSVQWNST 22  
Db 19 LLNLWGCKGRLVCYTSVKWNST 40

RESULT 15

US-09-444-410-37  
; Sequence 37, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/444,410  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/894,699  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-444-410-37

Query Match 85.8%; Score 109; DB 3; Length 40;  
Best Local Similarity 81.8%; Pred. No. 3.8e-09;  
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRLVCYTSVQWNST 22  
Db 19 LLNLWGCKGRLLCYTSVKWNST 40

Search completed: May 7, 2004, 17:53:25  
Job time : 12.1092 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 30.7815 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-5

Perfect score: 127

Sequence: 1 LLOSWGCKGRVCYTSVQWNST 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	87.4	113	14	US-10-320-786-40
2	111	87.4	715	14	Sequence 40, Appl
3	110	86.6	113	14	US-10-320-786-134
4	109	85.8	117	14	Sequence 134, Appl
5	108	85.0	116	14	Sequence 28, Appl
6	107	84.3	23	14	Sequence 6, Appl
7	107	84.3	113	14	US-10-320-786-20
8	107	84.3	113	14	Sequence 20, Appl
9	107	84.3	113	14	Sequence 137, Appl
10	107	84.3	215	8	Sequence 10, Appl
11	107	84.3	245	8	Sequence 12, Appl
12	107	84.3	281	8	US-08-911-824-58
13	107	84.3	373	8	Sequence 48, Appl
14	107	84.3	460	8	Sequence 58, Appl
15	107	84.3	490	8	US-08-911-824-120
					Sequence 120, Appl
					Sequence 52, Appl
					Sequence 60, Appl
					Sequence 95, Appl
					Sequence 50, Appl

16	107	84.3	526	8	US-08-911-824-97	Sequence 97, Appl
17	107	84.3	618	8	US-08-911-824-54	Sequence 54, Appl
18	107	84.3	706	8	US-08-911-824-93	Sequence 93, Appl
19	107	84.3	736	8	US-08-911-824-91	Sequence 91, Appl
20	107	84.3	873	8	US-08-911-824-61	Sequence 61, Appl
21	106	83.5	23	14	US-10-320-786-95	Sequence 95, Appl
22	106	83.5	110	14	US-10-320-786-14	Sequence 14, Appl
23	106	83.5	113	14	US-10-320-786-2	Sequence 2, Appl
24	106	83.5	113	14	US-10-320-786-18	Sequence 18, Appl
25	106	83.5	115	14	US-10-320-786-24	Sequence 24, Appl
26	105	82.7	110	14	US-10-320-786-16	Sequence 16, Appl
27	104	81.9	24	14	US-10-320-786-138	Sequence 138, Appl
28	104	81.9	37	14	US-10-026-741-86	Sequence 86, Appl
29	104	81.9	200	9	US-09-854-816-104	Sequence 104, Appl
30	104	81.9	356	14	US-10-357-400-12	Sequence 12, Appl
31	104	81.9	862	14	US-10-369-294-15	Sequence 15, Appl
32	103	81.1	24	14	US-10-320-786-92	Sequence 92, Appl
33	102	80.3	116	14	US-10-320-786-22	Sequence 22, Appl
34	101	79.5	23	14	US-10-320-786-91	Sequence 91, Appl
35	101	79.5	23	14	US-10-320-786-100	Sequence 99, Appl
36	101	79.5	23	14	US-10-320-786-101	Sequence 100, Appl
37	101	79.5	23	14	US-10-320-786-101	Sequence 101, Appl
38	101	79.5	35	14	US-10-026-741-101	Sequence 101, Appl
39	101	79.5	37	14	US-10-026-741-7	Sequence 7, Appl
40	101	79.5	37	14	US-10-026-741-90	Sequence 90, Appl
41	101	79.5	110	14	US-10-320-786-38	Sequence 38, Appl
42	101	79.5	113	14	US-10-320-786-4	Sequence 4, Appl
43	101	79.5	351	14	US-10-026-741-47	Sequence 47, Appl
44	101	79.5	877	14	US-10-026-741-102	Sequence 102, Appl
45	100	78.7	37	14	US-10-026-741-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1  
US-10-320-786-40  
; Sequence 40, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc\_feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-10-320-786-40

Query Match 87.4%; Score 111; DB 14; Length 113;  
Best Local Similarity 81.8%; Pred. No. 1.4e-07;  
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLOSWGCKGRVCYTSVQWNST 22  
|||  
Db 39 LLNLWGCKGRVCYTSVQWNKT 60



; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 137  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Human  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(23)  
; OTHER INFORMATION:  
US-10-320-786-137

Query Match 84.3%; Score 107; DB 14; Length 23;  
Best Local Similarity 90.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWN 20  
DB 4 LLNLWGCKGRLVCYTSVQWN 23

RESULT 7  
US-10-320-786-10  
; Sequence 10, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-10

Query Match 84.3%; Score 107; DB 14; Length 113;  
Best Local Similarity 77.3%; Pred. No. 4.9e-07;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22  
DB 39 LLNLWGCKGRLVCYTSVKWNNT 60

RESULT 8  
US-10-320-786-12  
; Sequence 12, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01

; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-12

Query Match 84.3%; Score 107; DB 14; Length 113;  
Best Local Similarity 77.3%; Pred. No. 4.9e-07;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22  
DB 39 LLNLWGCKGRLVCYTSVKWNNT 60

RESULT 9  
US-08-911-824-58  
; Sequence 58, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
; FILE REFERENCE: 6165.US.O1  
; CURRENT APPLICATION NUMBER: US/08/911,824  
; CURRENT FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
; FEATURE:  
; OTHER INFORMATION: Encodes recombinant protein pGO-8PL  
US-08-911-824-58

Query Match 84.3%; Score 107; DB 8; Length 215;  
Best Local Similarity 77.3%; Pred. No. 9e-07;  
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22  
DB 127 LLNLWGCKGRLVCYTSVKWNNT 148

RESULT 10  
US-08-911-824-48  
; Sequence 48, Application US/08911824  
; Publication No. US20030004323A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Hackett, John R., Jr.  
; APPLICANT: Yamaguchi, Julie  
; APPLICANT: Golden, Alan M.  
; APPLICANT: Brennan, Catherine A.  
; APPLICANT: Hickman, Robert K.  
; APPLICANT: Devare, Sushil G.  
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
; DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV

```
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9PL
US-08-911-824-48
```

```
Query Match      84.3%; Score 107; DB 8; Length 245;
Best Local Similarity 77.3%; Pred. No. 1e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LLQSWGCKGRILVCYTSVQWNST 22
   |||||||:||||:|
Db 127 LLNLWGCKGRILCYTSVKWNET 148
```

```
RESULT 11
US-08-911-824-120
; Sequence 120, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 120
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-15PL
US-08-911-824-120
```

```
Query Match      84.3%; Score 107; DB 8; Length 281;
Best Local Similarity 77.3%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LLQSWGCKGRILVCYTSVQWNST 22
   |||||||:||||:|
Db 127 LLNLWGCKGRILCYTSVKWNET 148
```

```
RESULT 12
US-08-911-824-52
; Sequence 52, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
```

```
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-11PL
US-08-911-824-52
```

```
Query Match      84.3%; Score 107; DB 8; Length 373;
Best Local Similarity 77.3%; Pred. No. 1.5e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LLQSWGCKGRILVCYTSVQWNST 22
   |||||||:||||:|
Db 127 LLNLWGCKGRILCYTSVKWNET 148
```

```
RESULT 13
US-08-911-824-60
; Sequence 60, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-8CKS
US-08-911-824-60
```

```
Query Match      84.3%; Score 107; DB 8; Length 460;
Best Local Similarity 77.3%; Pred. No. 1.8e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LLQSWGCKGRILVCYTSVQWNST 22
   |||||||:||||:|
Db 372 LLNLWGCKGRILCYTSVKWNET 393
```

```
RESULT 14
US-08-911-824-95
; Sequence 95, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
```

```
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-14PL
US-08-911-824-95
```

```
Query Match      84.3%; Score 107; DB 8; Length 488;
Best Local Similarity 77.3%; Pred. No. 1.9e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 LQSWGCKGRLCVYTSVQWNST 22
      ||| |||||:|||||:|
Db      127 LLNLWGCKGRLCYTSVKWNET 148
```

```
RESULT 15
US-08-911-824-50
; Sequence 50, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE
; FILE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08/911,824
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; OTHER INFORMATION: Encodes recombinant protein pGO-9CKS
US-08-911-824-50
```

```
Query Match      84.3%; Score 107; DB 8; Length 490;
Best Local Similarity 77.3%; Pred. No. 1.9e-06;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 LQSWGCKGRLCVYTSVQWNST 22
      ||| |||||:|||||:|
Db      372 LLNLWGCKGRLCYTSVKWNET 393
```

Search completed: May 7, 2004, 18:29:20  
Job time : 30.7815 secs

TABLE 31/11 (cont)



```
Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LQSWGCKGRLLVCYTSVQWNST 22
  || |||| ||:||||||:||
Db 607 LLNLWGCKRLLCYTSVKWNKT 628

RESULT 3
S22930
Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Jun-1995 #sequence_revision 21-Jul-1995 #text_change 26-Aug-1999
C:Accession: S52930
R:Cohen, J.H.M.; Guetard, D.; Philbert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de
submitted to the EMBL Data Library, January 1995
A:Description: A novel HIV1-O strain illustrates the diversity of the O group.
A:Reference number: S52929
A:Accession: S52930
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-104 <COH>
A:Cross-references: EMBL:X84328; NID:G695526; PIDN:CAA59066.1; PID:G695527
C:Superfamily: type E retrovirus env polyprotein

Query Match 78.7%; Score 100; DB 2; Length 104;
Best Local Similarity 68.2%; Pred. No. 2.2e-07; Indels 0; Gaps 0;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LQSWGCKGRLLVCYTSVQWNST 22
  || |||| ||:||||||:||
Db 45 LLNLWGCKRKAICYTSVQWNST 66

RESULT 4
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 20
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21990; S70423
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61357; NID:G60175; PIDN:CAA43626.1; PID:G60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70423
A:Accession: S70423
A:Molecule type: DNA
A:Residues: 1-332, X', 334-357 <STE2>
A:Cross-references: EMBL:X61357; NID:G60175; PIDN:CAA43626.1; PID:G60176
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 357;
Best Local Similarity 59.1%; Pred. No. 0.00018;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQSWGCKGRLLVCYTSVQWNST 22
  || |||| ||:||||||:||
Db 93 LLGIWGCSGRLLCTTAVPWNAS 114

RESULT 5
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 3L
```

```
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <STEL>
A:Cross-references: EMBL:X61352; NID:G60186; PIDN:CAA43616.1; PID:G60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, X', 335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:G60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 65.4%; Score 83; DB 2; Length 358;
Best Local Similarity 59.1%; Pred. No. 0.00018;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQSWGCKGRLLVCYTSVQWNST 22
  || |||| ||:||||||:||
Db 94 LLGIWGCSGRLLCTTAVPWNAS 115

RESULT 6
C41621
env polyprotein P - human immunodeficiency virus type 1 (fragment)
N:Alternate names: coat polyprotein
N:Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120
C:Species: human immunodeficiency virus type 1, HIV-1
A>Note: host Homo sapiens (man)
C:Date: 11-Feb-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C:Accession: C41621
R:Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991
A>Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A:Reference number: A41621; MUID:92107924; PMID:1763038
A:Accession: C41621
A:Molecule type: DNA
A:Residues: 1-443 <BUR>
A:Cross-references: GB:M77230; NID:G328631; PIDN:AAB03792.1; PID:G555015
A>Note: this virus was isolated from the mother's sexual partner
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>
F:252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>
F:424-443/Domain: transmembrane #status predicted <TMN>
F:9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: ca

Query Match 65.4%; Score 83; DB 2; Length 443;
Best Local Similarity 59.1%; Pred. No. 0.00022;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQSWGCKGRLLVCYTSVQWNST 22
  || |||| ||:||||||:||
Db 332 LLGIWGCSGRLLCTTAVPWNSS 353

RESULT 7
S54384
envelope polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C:Accession: S54384
R:Theodore, T.; Buckler-White, A.J.
```





```
S22006
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70420; S22006
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <STE2>
A:Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191
A:Experimental source: patient L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 54.5%; Pred. No. 0.00049;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRILVCYTSVQVNST 22
Db 93 LLGIWCGSGKLICTTAVPWNTS 114

RESULT 13
S21994
envelope protein gp120/gp41 - human immunodeficiency virus type 1
A:Variety: isolate 27B
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21994; S70421
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S21990
A:Accession: S21994
A:Molecule type: DNA
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70421
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140,'X',142-312,'X',314-357 <STE2>
A:Cross-references: EMBL:X61355; NID:g60179
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 54.5%; Pred. No. 0.00049;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRILVCYTSVQVNST 22
Db 93 LLGIWCGSGKLICTTAVPWNTS 114

RESULT 14
S22004
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 4B
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22004; S70419
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S21990
```

```
A:Accession: S22004
A:Molecule type: DNA
A:Residues: 1-357 <STE1>
A:Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292,'X',294-357 <STE2>
A:Cross-references: EMBL:X61353; NID:g60188
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 54.5%; Pred. No. 0.00049;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRILVCYTSVQVNST 22
Db 93 LLGIWCGSGKLICTTAVPWNTS 114

RESULT 15
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70422; S21996
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by polymerase chain reaction
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43624.1; PID:g1067129
A:Experimental source: patient 27L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match          63.0%; Score 80; DB 2; Length 357;
Best Local Similarity 54.5%; Pred. No. 0.00049;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRILVCYTSVQVNST 22
Db 93 LLGIWCGSGKLICTTAVPWNTS 114

Search completed: May 7, 2004, 17:51:02
Job time : 10.2437 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 5.54622 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-5  
Perfect score: 127  
Sequence: 1 LIQSWGCKGRLVCYTSVQWNST 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	65.4	853	1 ENV_HV122	P12487 human immun
2	83	65.4	854	1 ENV_SIVCZ	P17281 chimpanzee
3	83	65.4	855	1 ENV_HV126	P04580 human immun
4	80	63.0	843	1 ENV_HV1Y2	P35961 human immun
5	80	63.0	847	1 ENV_HV1S1	P19550 human immun
6	80	63.0	847	1 ENV_HV1W2	P05880 human immun
7	80	63.0	851	1 ENV_HV1B8	P04582 human immun
8	80	63.0	852	1 ENV_HV1B3	P12488 human immun
9	80	63.0	852	1 ENV_HV1S3	P19549 human immun
10	80	63.0	853	1 ENV_HV1MF	P19551 human immun
11	80	63.0	855	1 ENV_HV1A2	P03378 human immun
12	80	63.0	855	1 ENV_HV1OY	P20888 human immun
13	80	63.0	856	1 ENV_HV1B1	P03375 human immun
14	80	63.0	856	1 ENV_HV1H2	P04578 human immun
15	80	63.0	856	1 ENV_HV1LW	P06226 human immun
16	80	63.0	856	1 ENV_HV1MN	P05877 human immun
17	80	63.0	856	1 ENV_HV1PV	P03376 human immun
18	80	63.0	856	1 ENV_HV1SC	P05878 human immun
19	80	63.0	856	1 ENV_HV1W1	P31872 human immun
20	80	63.0	856	1 ENV_HV1ZH	P05881 human immun
21	80	63.0	861	1 ENV_HV1BR	P03377 human immun
22	80	63.0	865	1 ENV_HV1RH	P04579 human immun
23	80	63.0	867	1 ENV_HV1J3	P12489 human immun
24	80	63.0	868	1 ENV_HV1C4	P05879 human immun
25	79	62.2	846	1 ENV_HV1ND	P18799 human immun
26	78	61.4	848	1 ENV_HV1JR	P20871 human immun
27	78	61.4	856	1 ENV_HV1H3	P04624 human immun
28	76	59.8	821	1 ENV_SIVGB	P22380 simian immu
29	76	59.8	853	1 ENV_HV1EL	P04581 human immun
30	76	59.8	863	1 ENV_HV1Z8	P05882 human immun
31	74	58.3	859	1 ENV_HV1MA	P04583 human immun
32	74	58.3	861	1 ENV_HV1KB	P31819 human immun
33	73	57.5	857	1 ENV_HV2KR	Q74126 human immun

RESULT 1				
ENV_HV122	ENV_HV122	STANDARD;	PRT;	853 AA.
ID	AC	P12487;		
DT	01-OCT-1989	(Rel. 12, Created)		
DT	01-OCT-1989	(Rel. 12, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE	glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (23/CDC-234 isolate) (HIV-1).			
OC	Viruses; Retrovirdae; Retroviridae; Lentivirus.			
OX	NCBI_taxid=11683;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Theodore T., Buckler-White A.;			
RL	Submitted (NOV-1988) to the HIV data bank.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
EMBL;	M22639;	AAA45370.1;		
PIR;	S54384;	S54384.		
HIV;	M22639;	ENV52226.		
InterPro;	IPR000328;	Env_GP41.		
InterPro;	IPR000777;	GP120.		
Pfam;	PF00516;	GP120; 1.		
AIDS;	Coat protein;	Polyprotein; Glycoprotein; Transmembrane;		
Signal.				
SIGNAL	1	31	BY SIMILARITY.	
FT CHAIN	32	508	EXTERIOR MEMBRANE GLYCOPROTEIN.	
FT CHAIN	509	853	TRANSMEMBRANE GLYCOPROTEIN.	
FT DISULFID	53	73	BY SIMILARITY.	
FT DISULFID	118	206	BY SIMILARITY.	
FT DISULFID	125	197	BY SIMILARITY.	
FT DISULFID	130	154	BY SIMILARITY.	
FT DISULFID	219	248	BY SIMILARITY.	
FT DISULFID	229	240	BY SIMILARITY.	
FT DISULFID	297	330	BY SIMILARITY.	
FT DISULFID	376	442	BY SIMILARITY.	
FT DISULFID	383	415	BY SIMILARITY.	
FT CARBOHYD	87	87	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	137	137	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	144	144	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	153	153	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	157	157	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	185	185	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	188	188	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	198	198	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	235	235	N-LINKED (GLCNAC. . .)	(POTENTIAL).

34	71.5	56.3	856	1	ENV_HV2NZ	P05883 human immun
35	69	54.3	865	1	ENV_SIVAT	P05886 simian immu
36	68.5	53.9	712	1	ENV_HV2S2	P32536 human immun
37	68.5	53.9	846	1	ENV_HV2SB	P12449 human immun
38	68.5	53.9	859	1	ENV_HV2D2	P15831 human immun
39	68.5	53.9	859	1	ENV_HV2ST	P20872 human immun
40	68.5	53.9	885	1	ENV_SIVS4	P12492 simian immu
41	68	53.5	851	1	ENV_HV2D1	P17755 human immun
42	68	53.5	851	1	ENV_HV2G1	P18040 human immun
43	68	53.5	858	1	ENV_HV2RO	P04577 human immun
44	68	53.5	859	1	ENV_HV2CA	P24105 human immun
45	68	53.5	860	1	ENV_HV2BE	P18094 human immun

ALIGNMENTS



DR PIR; D26192; VCLJZR.  
DR HIV; K03458; ENV\$26.  
DR InterPro; IPR000328; Env GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Transmembrane;  
Signal.  
KW SIGNAL 1 30  
FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 207 BY SIMILARITY.  
FT DISULFID 125 198 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 220 249 BY SIMILARITY.  
FT DISULFID 230 241 BY SIMILARITY.  
FT DISULFID 298 332 BY SIMILARITY.  
FT DISULFID 378 444 BY SIMILARITY.  
FT DISULFID 385 417 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E233C3457 CRC64;  
Query Match 65.4%; Score 83; DB 1; Length 855;  
Best Local Similarity 59.1%; Pred. No. 4e-05;  
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LLQSGCKGRVLCYTSVQWNST 22  
DB 591 LLGIWCGSKGLICTTTVPWNS 612  
RESULT 4  
ID ENV\_HV1Y2 STANDARD; PRT; 843 AA.  
AC F35961;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polypeptide GP150 precursor (Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
GN ENV.

OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
NCBI\_TaxID=36377;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93021387; PubMed=1404605;  
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,  
Shaw G.M.;  
RT "Complete nucleotide sequence, genome organization, and biological  
properties of human immunodeficiency virus type 1 in vivo: evidence  
for limited defectiveness and complementation.";  
J. Virol. 66:6587-6600(1992).  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M93258; -; NOT ANNOTATED\_CDS.  
DR PIR; H44001; H44001.  
DR PDB; 1G9N; 27-DEC-00.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal;  
3D-structure. 1 29  
FT SIGNAL 1 29  
FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.  
FT TRANSMEM 738 755 POTENTIAL.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 201 BY SIMILARITY.  
FT DISULFID 125 192 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 214 243 BY SIMILARITY.  
FT DISULFID 224 235 BY SIMILARITY.  
FT DISULFID 292 326 BY SIMILARITY.  
FT DISULFID 373 432 BY SIMILARITY.  
FT DISULFID 380 405 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 843;  
Best Local Similarity 54.5%; Pred. No. 0.00011;  
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLOSWGCKGRLVCVTSVOWNST 22  
||| ||| :|:|:|:|:|:|:  
579 LLGIWGCSGKLICTTVPWNIS 600

Db

RESULT 5  
ENV\_HVI1S1 STANDARD; PRT; 847 AA.  
ID \_ENV\_HVI1S1 STANDARD; PRT; 847 AA.  
AC FL9550;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
EN ENV.  
GS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).  
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11691;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90347835; PubMed=2394920;  
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;  
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";  
RJ J. Virol. 64:4390-4398(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M65024; AAA45072.1; --  
CC HIV; LOBE; 15-MAY-97  
CC DRV; M38428; ENV\$SF162.  
CC DR InterPro; IPR000328; Env\_GP41.  
CC DR InterPro; IPR000777; GP120.  
CC DR Pfam; PF00516; GP120; 1.  
CC DR Pfam; PF00517; GP41; 1.  
CC KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.  
KW SIGNAL 1 29  
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.  
FT FT DISULFID 53 73 BY SIMILARITY.  
FT FT DISULFID 118 203 BY SIMILARITY.  
FT FT DISULFID 125 194 BY SIMILARITY.  
FT FT DISULFID 130 155 BY SIMILARITY.  
FT FT DISULFID 216 245 BY SIMILARITY.  
FT FT DISULFID 226 237 BY SIMILARITY.  
FT FT DISULFID 294 328 BY SIMILARITY.  
FT FT DISULFID 374 435 BY SIMILARITY.  
FT FT DISULFID 381 408 BY SIMILARITY.  
FT FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 63.0%; Score 80; DB 1; Length 847;  
Best Local Similarity 54.5%; Pred. No. 0.00011;  
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLOSWGCKGRLVCVTSVOWNST 22  
||| ||| :|:|:|:|:|:|:  
583 LLGIWGCSGKLICTTAVPNAS 604

Db

RESULT 6  
ENV\_HVIW2 STANDARD; PRT; 847 AA.  
ID \_ENV\_HVIW2 STANDARD; PRT; 847 AA.  
AC P05880;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
EN ENV.  
GS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11705;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86235450; PubMed=3012778;  
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D., Salahuddin S.Z., Wong-Straal F., Gallo R.C., Parks E.S., Parks W.P.;  
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS.";  
RJ Science 232:1548-1553(1986).  
CC CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M12507; AAB12990.1; --  
CC HIV; M12507; ENV\$WMJ2.  
CC DR InterPro; IPR000328; Env\_GP41.  
CC DR InterPro; IPR000777; GP120.  
CC DR Pfam; PF00516; GP120; 1.  
CC DR Pfam; PF00517; GP41; 1.  
CC KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.  
KW SIGNAL 1 29  
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.  
FT FT DISULFID 53 73 BY SIMILARITY.  
FT FT DISULFID 118 202 BY SIMILARITY.  
FT FT DISULFID 125 193 BY SIMILARITY.  
FT FT DISULFID 130 152 BY SIMILARITY.  
FT FT DISULFID 215 244 BY SIMILARITY.  
FT FT DISULFID 225 236 BY SIMILARITY.  
FT FT DISULFID 293 326 BY SIMILARITY.





DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11693;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89085613; PubMed=2789516;  
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,  
RA Dandekar S.;  
RT "Biological and molecular characterization of human immunodeficiency  
RT virus (HIV-1BR) from the brain of a patient with progressive  
RT dementia.";  
RL Virology 168:79-89 (1989).  
CC -!- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS  
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M21098; AAA44221.1; -.  
DR PIR; A31667; VCLJBR.  
DR PDB; 1IM7; 23-OCT-02.  
DR HIV; M21098; ENV5BRVA.  
DR InterPro; IPR000328; Env GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
KW 3D-structure.  
FT SIGNAL 1 30  
FT CHAIN 31 507 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 205 BY SIMILARITY.  
FT DISULFID 126 196 BY SIMILARITY.  
FT DISULFID 131 155 BY SIMILARITY.  
FT DISULFID 218 247 BY SIMILARITY.  
FT DISULFID 228 239 BY SIMILARITY.  
FT DISULFID 236 330 BY SIMILARITY.  
FT DISULFID 376 439 BY SIMILARITY.  
FT DISULFID 383 412 BY SIMILARITY.  
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;  
Query Match 63.0%; Score 80; DB 1; Length 852;  
Best Local Similarity 54.5%; Pred. No. 0.00011;  
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LQSWGCKGRVLCYTSVQWNST 22  
Db 588 LLGIWGCCKLICCTTAVPWNAS 609  
RESULT 9  
ENV HV1S3  
ID ENV HV1S3 STANDARD; PRT; 852 AA.  
AC P19549;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90317906; PubMed=2370688;  
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;  
RT "Human immunodeficiency virus type 1 cellular host range,  
RT replication, and cytopathicity are linked to the envelope region of  
RT the viral genome.";  
RL J. Virol. 64:4016-4020 (1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AY352275; AAQ17031.1; -.  
DR PDB; 1MEQ; 11-DEC-02.  
DR HIV; M38427; ENV5SF33.  
DR InterPro; IPR000328; Env GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
KW 3D-structure.  
FT SIGNAL 1 31 BY SIMILARITY.  
FT CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 507 852 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 206 BY SIMILARITY.  
FT DISULFID 125 197 BY SIMILARITY.  
FT DISULFID 130 156 BY SIMILARITY.  
FT DISULFID 219 248 BY SIMILARITY.  
FT DISULFID 229 240 BY SIMILARITY.  
FT DISULFID 297 331 BY SIMILARITY.  
FT DISULFID 377 439 BY SIMILARITY.  
FT DISULFID 384 412 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).





-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; K02007; A059882.1; -;  
PIR; A03976; VCLJ2A.  
HIV; K02007; ENV5SF2.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
KW SIGNAL.  
FT CHAIN 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 208 BY SIMILARITY.  
FT DISULFID 125 199 BY SIMILARITY.  
FT DISULFID 130 155 BY SIMILARITY.  
FT DISULFID 221 250 BY SIMILARITY.  
FT DISULFID 231 242 BY SIMILARITY.  
FT DISULFID 239 333 BY SIMILARITY.  
FT DISULFID 380 442 BY SIMILARITY.  
FT DISULFID 387 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 855;  
Best Local Similarity 54.5%; Pred. No. 0.00011;  
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LIQSGKCKGRLYCYTSVQWNST 22

Db 591 LUGWGCSGLKICTTAVPWNAS 612

RESULT 12

ENV\_HV10Y  
ID ENV\_HV10Y STANDARD; PRT; 855 AA.  
AC P20888;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90148544; PubMed=2559749;  
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;  
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese  
individual presenting an atypical western blot.";  
RL AIDS 3:707-715(1989).  
CC -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A  
HEALTHY GABONESE INDIVIDUAL.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M26727; AAA83397.1; -;  
HIV; M26727; ENV5OYI.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
Signal.  
KW SIGNAL.  
FT CHAIN 1 29  
FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 210 BY SIMILARITY.  
FT DISULFID 125 201 BY SIMILARITY.  
FT DISULFID 130 162 BY SIMILARITY.  
FT DISULFID 223 252 BY SIMILARITY.  
FT DISULFID 233 244 BY SIMILARITY.  
FT DISULFID 301 335 BY SIMILARITY.  
FT DISULFID 381 442 BY SIMILARITY.  
FT DISULFID 388 415 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 63.0%; Score 80; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.00011;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22
Db 591 LLGIWCGSKLICITTVAPWNAS 612

RESULT 13
ENV HV1B1
ID ENV HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RT Nature 313:277-284 (1985).
RL Nature 313:277-284 (1985).
RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RA "Assessment of intrachain disulfide bonds and characterization of
potential glycosylation sites of the type 1 recombinant human
immunodeficiency virus envelope glycoprotein (gp120) expressed in
Chinese hamster ovary cells";
RT J. Biol. Chem. 265:10373-10382 (1990).
RL J. Biol. Chem. 265:10373-10382 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15654; AAA44205.1; --
DR PIR; A03973; VCLJH3.
DR HIV; M15654; ENV5BH102.
DR InterPro; IPR000328; Env.GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
KW SIGNAL 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157

```

```

FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .)
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .)
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .)
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .)
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .)
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .)
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .)
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 856;
Best Local Similarity 54.5%; Pred. No. 0.00012;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22
Db 592 LLGIWCGSKLICITTVAPWNAS 613

RESULT 14
ENV HV1H2
ID ENV HV1H2 STANDARD; PRT; 856 AA.
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RA "Complete nucleotide sequences of functional clones of the AIDS
virus.";
RT AIDS Res. Hum. Retroviruses 3:57-69 (1987).
RL AIDS Res. Hum. Retroviruses 3:57-69 (1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

```





COPIES OF THE

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 27.916 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362a-5  
Perfect score: 127  
Sequence: 1 LLQSWGCKGRLVCYTSVQWNST 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	90.6	219	15 Q9IEB6	Q9ieb6 human immun
2	114	89.8	116	15 Q7ZJN9	Q7zjn9 human immun
3	114	89.8	135	15 Q9DQL9	Q9dql9 human immun
4	114	89.8	216	15 Q9IEC5	Q9iec5 human immun
5	114	89.8	242	15 Q9IE31	Q9ie31 human immun
6	112	88.2	214	15 Q9DIK3	Q9dik3 human immun
7	111	87.4	116	15 O11941	O11941 human immun
8	111	87.4	116	15 O40458	O40458 human immun
9	111	87.4	130	15 Q9IHU9	Q9ihu9 human immun
10	111	87.4	158	15 Q8J3N6	Q8j3n6 human immun
11	111	87.4	183	15 Q9IEC1	Q9iec1 human immun
12	111	87.4	218	15 Q9IE95	Q9ie95 human immun
13	111	87.4	220	15 Q9IEC9	Q9iec9 human immun
14	111	87.4	512	15 Q9IED2	Q9ied2 human immun
15	111	87.4	538	15 Q9IED5	Q9ied5 human immun
16	111	87.4	872	15 Q8Q7H0	Q8q7h0 human immun

17	111	87.4	879	15 Q9WIU9	Q9wiu9 human immun
18	111	87.4	880	15 Q8Q7I9	Q8q7i9 human immun
19	111	87.4	882	15 Q8Q7F9	Q8q7f9 human immun
20	111	87.4	887	15 Q8Q7H6	Q8q7h6 human immun
21	111	87.4	887	15 Q8Q7G9	Q8q7g9 human immun
22	111	87.4	900	15 Q9QNZ8	Q9qnz8 human immun
23	110	86.6	114	15 O40448	O40448 human immun
24	110	86.6	235	15 Q9IE54	Q9ie54 human immun
25	109	85.8	118	15 O40451	O40451 human immun
26	109	85.8	214	15 Q9IE96	Q9ie96 human immun
27	109	85.8	216	15 Q9IEC7	Q9iec7 human immun
28	109	85.8	216	15 Q9IEA5	Q9iea5 human immun
29	109	85.8	219	15 Q9IEC8	Q9iec8 human immun
30	109	85.8	234	15 Q9IEA6	Q9iea6 human immun
31	109	85.8	242	15 Q9IE30	Q9ie30 human immun
32	109	85.8	535	15 Q9IEF2	Q9ief2 human immun
33	109	85.8	544	15 Q9IED9	Q9ied9 human immun
34	109	85.8	548	15 Q9IED6	Q9ied6 human immun
35	109	85.8	551	15 Q9IEE1	Q9iee1 human immun
36	109	85.8	865	15 Q8Q7H7	Q8q7h7 human immun
37	109	85.8	867	15 Q8Q7G8	Q8q7g8 human immun
38	108	85.0	116	15 O40459	O40459 human immun
39	108	85.0	134	15 Q9IHV4	Q9ihv4 human immun
40	108	85.0	137	15 Q9IHV5	Q9ihv5 human immun
41	108	85.0	146	15 Q9WRV2	Q9wrv2 human immun
42	108	85.0	200	15 Q9IEB8	Q9ieb8 human immun
43	108	85.0	227	15 Q9IE99	Q9ie99 human immun
44	108	85.0	238	15 Q9DIK1	Q9dik1 human immun
45	108	85.0	517	15 Q9IEE7	Q9iee7 human immun

ALIGNMENTS

RESULT 1

Q9IEB6 PRELIMINARY; PRT; 219 AA.

AC Q9IEB6; (1)

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Gp41 (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BCF14;

RA Roques P., Robertson D., Sandrine S., Christel D., Francois S., Philippe M.;

RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ236404; CAB96252.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR000328; Env\_GP41.

DR Pfam; PF00517; GP41; 1.

DR KW Transmembrane.

FT NON\_TER 1

FT NON\_TER 219

SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 90.6%; Score 115; DB 15; Length 219;  
Best Local Similarity 86.4%; Pred. No. 3.9e-10;  
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLQSWGCKGRLVCYTSVQWNST 22  
Db 57 LLNSWGCKGRLVCYTSVRKWNST 78

```
RESULT 2
Q7ZJN9          PRELIMINARY;      PRT;    116 AA.
ID Q7ZJN9
AC Q7ZJN9
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08692A;
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol
RT integrase, and env gp41.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY214120; AAO61840.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match      89.8%; Score 114; DB 15; Length 116;
Best Local Similarity 86.4%; Pred. No. 3e-10;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22
DB 32 LQSWGCKGRLVCYTSVQWNST 53

RESULT 3
Q9DQL9          PRELIMINARY;      PRT;    135 AA.
ID Q9DQL9
AC Q9DQL9
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20384646; PubMed=11153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,
RA Vadiello J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).
DR EMBL; AF255939; AAG36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 135 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match      89.8%; Score 114; DB 15; Length 135;
Best Local Similarity 81.8%; Pred. No. 3.5e-10;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22
```

```
Db 48 LQSWGCKGRLVCYTSVQWNST 69

RESULT 4
Q9IEC5          PRELIMINARY;      PRT;    216 AA.
ID Q9IEC5
AC Q9IEC5
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF07;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236394; CAB96243.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match      89.8%; Score 114; DB 15; Length 216;
Best Local Similarity 86.4%; Pred. No. 5.6e-10;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQSWGCKGRLVCYTSVQWNST 22
DB 47 LQSWGCKGRLVCYTSVQWNST 68

RESULT 5
Q9IEJ1          PRELIMINARY;      PRT;    242 AA.
ID Q9IEJ1
AC Q9IEJ1
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF112;
RA Roques P., Robertson D., Sousquiere S., Damond F., Mauciere P.,
RA Delienne C., Brun-Vezinet F., Dormont D., Simon F.o.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243366; CAB96336.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 242 242
SQ SEQUENCE 242 AA; 27539 MW; DE8A73DF0E8A6FD7 CRC64;
```



```

DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane. 1
FT NON_TER 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 488A9A40F4255E3E CRC64;
Query Match 87.4%; Score 111; DB 15; Length 116;
Best Local Similarity 81.8%; Pred. No. 8.8e-10;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 LLSWGCKGRLCVYTSVQWNST 22
Db 39 LLNLWGCKGRLCVYTSVQWNST 60
RESULT 8
O04058 PRELIMINARY; PRT; 116 AA.
AC O04058;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Envelope transmembrane glycoprotein (Fragment).
DE Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=group O;
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;
RT "Molecular characterization of envelope transmembrane glycoprotein of
RT 14 new human immunodeficiency virus type 1 group O strains from
RT different African countries.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RM ENBL; Y03774; CAA70913.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane. 1
FT NON_TER 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 540F5B3B7A849D8 CRC64;
Query Match 87.4%; Score 111; DB 15; Length 116;
Best Local Similarity 81.8%; Pred. No. 8.8e-10;
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 LLSWGCKGRLCVYTSVQWNST 22
Db 39 LLNLWGCKGRLCVYTSVQWNKT 60
RESULT 9
Q9IHU9 PRELIMINARY; PRT; 130 AA.
AC Q9IHU9;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OC Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=109333623;

```

```
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229235; RAF71912.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 130
SQ SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match 87.4%; Score 111; DB 15; Length 130;
Best Local Similarity 81.8%; Pred. No. 9.9e-10;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 40 LLNSWGCKGKGLVCYTSVKWNKT 61
|| |||||:|||||:|||||

RESULT 10
ID Q8J3N6 PRELIMINARY; PRT; 158 AA.
AC Q8J3N6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 protein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22092513; PubMed=12097573;
RA Boutonnet N., Janssens W., Bouton C., Verschelde J.L., Heyndrickx L.,
RA Beirnaert E., van der Groen G., Lasters I.;
RT "Comparison of predicted scaffold-compatible sequence variation in the
RT triple-hairpin structure of human immunodeficiency virus type 1 gp41
RT with patient data.";
RL J. Virol. 76:7595-7606(2002).
DR EMBL; AJ428017; CAD20969.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 158
SQ SEQUENCE 158 AA; 18803 MW; 6B8AD85559B33C8 CRC64;

Query Match 87.4%; Score 111; DB 15; Length 158;
Best Local Similarity 81.8%; Pred. No. 1.2e-09;
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 48 LLNLWGCKGRLVCYTSVQWNKT 69
|| |||||:|||||:|||||

RESULT 11
ID Q9IEC1 PRELIMINARY; PRT; 183 AA.
AC Q9IEC1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```

```
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF101;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236398; CAB96247.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 183
SQ SEQUENCE 183 AA; 21722 MW; C22A28A73540C8AA CRC64;

Query Match 87.4%; Score 111; DB 15; Length 183;
Best Local Similarity 81.8%; Pred. No. 1.4e-09;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 42 LLNLWGCKGRLVCYTSVKWNST 63
|| |||||:|||||:|||||

RESULT 12
ID Q9IE95 PRELIMINARY; PRT; 218 AA.
AC Q9IE95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF51;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236425; CAB96273.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 218
SQ SEQUENCE 218 AA; 25243 MW; E7F0E1E20EF79FA8 CRC64;

Query Match 87.4%; Score 111; DB 15; Length 218;
Best Local Similarity 77.3%; Pred. No. 1.7e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTSVQWNST 22
DB 54 LLSSWGCKGRLVCYTSVKWNKT 75
|| |||||:|||||:|||||

RESULT 13
Q9IEC9
```

```

Qy 1 LLSWGCKGRLVCYTSVQWNST 22
   || |||||:||||:||||
Db 372 LLNLWGCKGLICYTSVKWNST 393

RESULT 15
QRIED5 PRELIMINARY; PRT; 538 AA.
AC QRIED5;
DT 01-OCT-2000- (TtEMBLrel. 15, Created)
DT 01-OCT-2000 (TtEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TtEMBLrel. 24, Last annotation update)
DE Env polypeptide (Fragment).
EN GN.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF51;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA Depienne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDJ databases.
DR EMBL; AJ133072; CAB96233.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 538 538
SQ SEQUENCE 538 AA; 60777 MW; B3C9E6A233FEF1D CRC64;

Query Match 87.4%; Score 111; DB 15; Length 538;
Best Local Similarity 77.3%; Pred. No. 4.1e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0;

Qy 1 LLSWGCKGRLVCYTSVQWNST 22
   || |||||:||||:||||
Db 374 LLSSWGCKGLICYTSVKWNST 395

Search completed: May 7, 2004, 17:49:12
Job time : 27.916 secs

```

SECRET

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 41.8739 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-8  
Perfect score: 127  
Sequence: 1 LLSWGCRLVCYTSVQWNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	100.0	22	AAW80466	Peptide d
2	124	97.6	22	AAW80461	Peptide d
3	121	95.3	22	AAW80460	Peptide d
4	120	94.5	32	AAW80469	Peptide d
5	119	93.7	22	AAW80462	Peptide d
6	117	92.1	40	AAW07346	Partial s
7	116	91.3	22	AAW80465	Peptide d
8	116	91.3	22	AAW80464	Peptide d
9	115	90.6	22	AAW80463	Peptide d
10	114	89.8	32	AAW80470	Peptide d
11	113	89.0	22	AAW80459	Peptide d
12	113	89.0	32	AAW80471	Peptide d
13	110	86.6	40	AAW07352	Partial s
14	110	86.6	113	AAAY05565	AAAY05565 HIV-1 gro
15	110	86.6	200	AAAY77373	HIV-1 gro
16	110	86.6	215	AAAY09499	HIV-1 gro
17	110	86.6	215	AAAY06983	Recombina
18	110	86.6	215	AAAY77374	HIV-1 gro
19	110	86.6	245	AAAY09493	HIV-1 gro
20	110	86.6	245	AAAY06977	Recombina
21	110	86.6	245	AAAY77369	HIV-1 gro
22	110	86.6	281	AAAY09507	HIV-1 gro
23	110	86.6	373	AAAY09495	HIV-1 gro
24	110	86.6	373	AAAY06979	Recombina
25	110	86.6	460	AAAY09500	HIV-1 gro

## ALIGNMENTS

RESULT 1

AAW80466

ID AAW80466 standard; peptide; 22 AA.

XX AAW80466;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)

XX Peptide derived from a conserved sequence of group O human HIV.

XX Group O human immune deficiency virus; HIV; detection; infection.

XX Synthetic.

OS Human immunodeficiency virus 1.

XX WO9845323-A1.

XX 15-OCT-1998.

XX 06-APR-1998; 98WO-FR000691.

XX 09-APR-1997; 97FR-00004356.

XX 24-FEB-1998; 98FR-00002212.

XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX Cheneboux DMB, Delagneau JH, Gadelle SUX, Rieunier PY;

XX WPI; 1998-583190/49.

XX New synthetic peptide(s) - useful for, e.g. detecting infection by human immune deficiency virus of group O.

XX Claim 6; Page 43; 55pp; French.

XX AAW80459-74 represent synthetic peptides (either linear or cyclised by Cys-Cys disulphide bonds). The peptides represent variable sequences connected around short highly conserved sequences present in isolates of group O human immune deficiency virus (HIV). The peptides are useful as immunological reagents for detecting infection by group O human immune deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 22 AA;

Query Match 100.0%; Score 127; DB 2; Length 22;

26 110 86.6 460 2 AAY06984 Recombina  
27 110 86.6 460 3 AAY77375  
28 110 86.6 474 3 AAY77371  
29 110 86.6 488 2 AAY09504  
30 110 86.6 490 2 AAY09494 HIV-1 Gro  
31 110 86.6 490 2 AAY06978 Recombina  
32 110 86.6 490 3 AAY77370 HIV-1 gro  
33 110 86.6 526 2 AAY09505  
34 110 86.6 618 2 AAY09496  
35 110 86.6 618 2 AAY06980 Recombina  
36 110 86.6 618 3 AAY77372 HIV-1 gro  
37 110 86.6 706 2 AAY09503 HIV-1 gro  
38 110 86.6 715 2 AAY05625 HIV-1 gro  
39 110 86.6 736 2 AAY09502 HIV-1 gro  
40 110 86.6 873 2 AAY09501 HIV-1 gro  
41 110 86.6 873 2 AAY06985 Amino aci  
42 110 86.6 873 3 AAY77376 HIV-1 gro  
43 109 85.8 104 2 AAW07245 HIV-1 gro  
44 108 85.0 33 3 AAB12212 Partial s  
45 107 84.3 33 3 AAB12231 Partial s



PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gabelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX Claim 6; Page 44; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 32 AA;  
 Query Match 94.5%; Score 120; DB 2; Length 32;  
 Best Local Similarity 90.9%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWCGRGLVCYTSVQWNET 22  
 ||:|||||:|||||:|||||  
 Db 11 LLNSWCGRGLVCYTSVRWNET 32  
 ||:|||||:|||||:|||||  
 RESULT 5  
 AAW80462  
 ID AAW80462 standard; peptide; 22 AA.  
 XX  
 AC AAW80462;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 DE Group O human immune deficiency virus; HIV; detection; infection.  
 XX  
 KW Group O human immune deficiency virus 1.  
 OS Synthetic.  
 XX Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gabelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX Claim 6; Page 42; 55pp; French.  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC

CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 22 AA;  
 Query Match 93.7%; Score 119; DB 2; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 1.4e-08;  
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LLSSWCGRGLVCYTSVQWNET 22  
 ||:|||||:|||||:|||||  
 Db 1 LLSSWCGRGLVCYTSVQWNET 22  
 ||:|||||:|||||:|||||  
 RESULT 6  
 AAW07346  
 ID AAW07346 standard; peptide; 40 AA.  
 XX  
 AC AAW07346;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 03-JUN-1997 (first entry)  
 XX  
 DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).  
 XX  
 KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX WO9627013-A1.  
 XX  
 PD 06-SEP-1996.  
 XX  
 PF 26-FEB-1996; 96WO-FR000294.  
 XX  
 PR 27-FEB-1995; 95FR-00002236.  
 XX  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.  
 XX  
 PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;  
 XX  
 DR WPI; 1996-412779/41.  
 DR N-PSDB; AAT44922.  
 XX  
 PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and  
 PT antibodies - useful for diagnosis, screening and typing, or as  
 PT immunogens.  
 XX  
 PS Claim 12; Page 34; 71pp; French.  
 XX  
 CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided  
 CC into 2 major groups based on the nucleotide sequences of the envelop gene  
 CC (env): group M containing sub-groups A-G, and group O containing the  
 CC strains ANT70 and MVP180. The invention relates to the discovery of  
 CC several new strains of HIV-1 which can be placed in group O, based on the  
 CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39  
 CC and AAW07329-64). The novel strains have been deposited as retroviruses  
 CC CNCM I-1544 (BCF02 (BSS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547  
 CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from  
 CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein  
 CC encoded by the env gene. The nucleic acids can be used to detect gp. O  
 CC HIV-1 strains by hybridisation or (as primers) by gene amplification.  
 CC also for screening and typing of such strains. Peptides encoded by the  
 CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 40 AA;

```
Query Match          92.1%; Score 117; DB 2; Length 40;
Best Local Similarity 86.4%; Pred. No. 4.6e-08;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQWNET 22
Db 19 LLSSWGCRGLVCYTSVQWNET 40
      ||:||||:||||:||||:||||:
      ||:||||:||||:||||:||||:

RESULT 7
AAW80465
ID AAW80465 standard; peptide; 22 AA.
XX
AC AAW80465;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
OS Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
PI
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 43; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 22 AA;

Query Match          91.3%; Score 116; DB 2; Length 22;
Best Local Similarity 86.4%; Pred. No. 3.5e-08;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQWNET 22
Db 1 LLSSWGCRGLVCYTSVQWNET 22
      ||:||||:||||:||||:||||:
      ||:||||:||||:||||:||||:

RESULT 9
AAW80463
ID AAW80463 standard; peptide; 22 AA.
XX
XX AAW80463;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-JAN-1999 (first entry)
XX
DE Peptide derived from a conserved sequence of group O human HIV.
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
OS Human immunodeficiency virus 1.
XX
XX WO9845323-A1.
XX
XX 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-FR000691.
XX
XX 09-APR-1997; 97FR-00004356.
PR 24-FEB-1998; 98FR-00002212.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
PI
XX WPI; 1998-583190/49.
XX
XX New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX immune deficiency virus of group O.
XX
XX Claim 6; Page 43; 55pp; French.
XX
XX AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX Cys-Cys disulphide bonds). The peptides represent variable sequences
XX connected around short highly conserved sequences present in isolates of
XX group O human immune deficiency virus (HIV). The peptides are useful as
XX immunological reagents for detecting infection by group O human immune
XX deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 22 AA;

Query Match          91.3%; Score 116; DB 2; Length 22;
Best Local Similarity 90.9%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQWNET 22
Db 1 LLSSWGCRGLVCYTSVQWNET 22
      |||:||||:||||:||||:||||:
      |||:||||:||||:||||:||||:

RESULT 8
AAW80464
ID AAW80464 standard; peptide; 22 AA.
XX
XX AAW80464;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
```



XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX DR WPI; 1998-583190/49.  
 XX PT New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX PT immune deficiency virus of group O.  
 XX PS Claim 6; Page 42; 55pp; French.  
 XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 XX CC connected around short highly conserved sequences present in isolates of  
 XX CC group O human immune deficiency virus (HIV). The peptides are useful as  
 XX CC immunological reagents for detecting infection by group O human immune  
 XX CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 XX CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX SQ Sequence 22 AA;  
 Query Match 90.6%; Score 115; DB 2; Length 22;  
 Best Local Similarity 86.4%; Pred. No. 4.8e-08;  
 Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSWGCGRGLVCYTSVQWNST 22  
 |||:|||||:|||||  
 Db 1 LLQSWGCKGRGLVCYTSVQWNST 22  
 |||:|||||:|||||  
 RESULT 10  
 AAW80470  
 ID AAW80470 standard; peptide; 32 AA.  
 AC AAW80470;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 DE Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX PS Claim 6; Page 44; 55pp; French.  
 XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 XX CC connected around short highly conserved sequences present in isolates of  
 XX CC group O human immune deficiency virus (HIV). The peptides are useful as  
 XX CC immunological reagents for detecting infection by group O human immune

CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX SQ Sequence 32 AA;  
 Query Match 89.8%; Score 114; DB 2; Length 32;  
 Best Local Similarity 86.4%; Pred. No. 9.1e-08;  
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 LLSWGCGRGLVCYTSVQWNST 22  
 |||:|||||:|||||  
 Db 1 LLNWSGCRGLVCYTSVRWNET 32  
 |||:|||||:|||||  
 RESULT 11  
 AAW80459  
 ID AAW80459 standard; peptide; 22 AA.  
 XX AAW80459;  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX Peptide derived from a conserved sequence of group O human HIV.  
 DE Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9845323-A1.  
 XX 15-OCT-1998.  
 XX 06-APR-1998; 98WO-FR000691.  
 XX 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 XX immune deficiency virus of group O.  
 XX PS Claim 6; Page 42; 55pp; French.  
 XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 XX CC connected around short highly conserved sequences present in isolates of  
 XX CC group O human immune deficiency virus (HIV). The peptides are useful as  
 XX CC immunological reagents for detecting infection by group O human immune  
 XX CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 XX CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX SQ Sequence 22 AA;  
 Query Match 89.0%; Score 113; DB 2; Length 22;  
 Best Local Similarity 86.4%; Pred. No. 8.6e-08;  
 Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LLSWGCGRGLVCYTSVQWNST 22  
 |||:|||||:|||||  
 Db 1 LLSWGCGRGLVCYTSVQWNST 22  
 |||:|||||:|||||  
 RESULT 12  
 AAW80471  
 ID AAW80471 standard; peptide; 32 AA.





SECRET

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 12.1092 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-8  
Perfect score: 127  
Sequence: 1 LLSWCGRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	92.1	40	3	US-08-894-699-39
2	117	92.1	40	3	US-09-444-410-39
3	110	86.6	40	3	US-08-894-699-68
4	110	86.6	40	3	US-09-444-410-68
5	110	86.6	113	4	US-09-462-917A-40
6	110	86.6	215	2	US-08-912-129A-58
7	110	86.6	245	2	US-08-912-129A-48
8	110	86.6	373	2	US-08-912-129A-52
9	110	86.6	460	2	US-08-912-129A-60
10	110	86.6	490	2	US-08-912-129A-50
11	110	86.6	618	2	US-08-912-129A-54
12	110	86.6	715	4	US-09-462-917A-134
13	110	86.6	873	2	US-08-912-129A-61
14	109	85.8	37	4	US-08-817-441-94
15	109	85.8	104	4	US-08-817-441-100
16	108	85.0	33	3	US-09-433-428D-6
17	107	84.3	33	3	US-09-433-428D-25
18	107	84.3	40	3	US-08-894-699-36
19	107	84.3	40	3	US-09-444-410-36
20	107	84.3	113	4	US-09-462-917A-28
21	107	84.3	116	4	US-09-462-917A-20
22	107	84.3	356	1	US-08-602-713-12
23	107	84.3	356	3	US-08-989-493-12
24	107	84.3	356	4	US-09-610-271-12
25	106	83.5	37	4	US-08-817-441-86
26	106	83.5	41	3	US-08-894-699-69
27	106	83.5	41	3	US-09-444-410-69

28	106	83.5	113	4	US-09-462-917A-2	Sequence 2, Appli
29	106	83.5	113	4	US-09-462-917A-18	Sequence 18, Appli
30	106	83.5	117	4	US-09-462-917A-6	Sequence 6, Appli
31	106	83.5	200	3	US-08-965-056-104	Sequence 104, App
32	106	83.5	862	4	US-09-206-551-15	Sequence 15, Appl
33	105	82.7	23	4	US-08-817-441-30	Sequence 30, Appl
34	105	82.7	23	4	US-09-462-917A-137	Sequence 137, App
35	105	82.7	33	3	US-09-433-428D-30	Sequence 30, Appl
36	105	82.7	41	3	US-08-894-699-67	Sequence 67, Appl
37	105	82.7	41	3	US-09-444-410-67	Sequence 67, Appl
38	105	82.7	110	4	US-09-462-917A-14	Sequence 14, Appl
39	105	82.7	115	4	US-09-462-917A-24	Sequence 24, Appl
40	105	82.7	116	4	US-09-462-917A-22	Sequence 22, Appl
41	104	81.9	23	4	US-09-462-917A-95	Sequence 95, Appl
42	104	81.9	33	3	US-09-433-428D-8	Sequence 8, Appli
43	104	81.9	40	3	US-08-894-699-41	Sequence 41, Appl
44	104	81.9	40	3	US-08-894-699-42	Sequence 42, Appl
45	104	81.9	40	3	US-09-444-410-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1  
US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-894-699-39

Query Match 92.1%; Score 117; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 2.6e-10;  
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22  
DB 19 LLNSWGCKGRLVCYTSVKWNET 40

RESULT 2

US-09-444-410-39  
Sequence 39, Application US/09444410  
Patent No. 6270975

GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSET-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-09-444-410-39

Query Match 92.1%; Score 117; DB 3; Length 40;  
Best Local Similarity 86.4%; Pred. No. 2.6e-10;  
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22  
DB 19 LLNSWGCKGRLVCYTSVKWNET 40

RESULT 3

US-08-894-699-68  
Sequence 68, Application US/08894699  
Patent No. 6030769

GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSET-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,699  
FILING DATE: 01-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR96/00294  
FILING DATE: 26-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-894-699-68

Query Match 86.6%; Score 110; DB 3; Length 40;  
Best Local Similarity 81.8%; Pred. No. 2.7e-09;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22  
DB 19 LLNSWGCKGRLVCYTSVQWNET 40

RESULT 4

US-09-444-410-68  
Sequence 68, Application US/09444410  
Patent No. 6270975

GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSET-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE

;; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
;; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; STREET: FLOOR  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/444,410  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/894,699  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614  
;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 68:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 86.6%; Score 110; DB 3; Length 40;  
Best Local Similarity 81.8%; Pred. No. 2.7e-09;  
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQVNET 22  
||:||||:||||:||||:||||:  
Db 19 LLNSWGCRGLVCYTSVEVNW 40

RESULT 5  
US-09-462-917A-40  
;; Sequence 40, Application US/09462917A  
;; Patent No. 6511801  
;; GENERAL INFORMATION:  
;; APPLICANT: Delaporte, Eric  
;; APPLICANT: Peeters, Martine  
;; APPLICANT: Saman, Eric  
;; APPLICANT: Vanden Haesevelde, Marlen  
;; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
;; FILE REFERENCE: INNS.014 11362.0014.NPUS00  
;; CURRENT APPLICATION NUMBER: US/09/462,917A  
;; CURRENT FILING DATE: 2000-04-03  
;; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
;; PRIOR FILING DATE: 1998-07-20  
;; NUMBER OF SEQ ID NOS: 152  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 40  
;; LENGTH: 113  
;; TYPE: PRT  
;; ORGANISM: Human

;; FEATURE:  
;; NAME/KEY: Misc\_feature  
;; LOCATION: (1)..(113)  
;; OTHER INFORMATION: Xaa = unknown  
US-09-462-917A-40  
Query Match 86.6%; Score 110; DB 4; Length 113;  
Best Local Similarity 77.3%; Pred. No. 7.9e-09;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQVNET 22  
||:||||:||||:||||:||||:  
Db 39 LLNLWGCKRGLICYTSVQWNKT 60

RESULT 6  
US-08-912-129A-58  
;; Sequence 58, Application US/08912129A  
;; Patent No. 5922533  
;; GENERAL INFORMATION:  
;; APPLICANT: VALLARI, ANADRUZELA S.  
;; APPLICANT: HACKETT, JOHN JR.  
;; APPLICANT: HICKMAN, ROBERT K.  
;; APPLICANT: VARITEK, VINCENT A. JR.  
;; APPLICANT: NECKLAWS, ELIZABETH A.  
;; APPLICANT: GOLDEN, ALAN M.  
;; APPLICANT: BRENNAN, CATHERINE A.  
;; APPLICANT: DEVARE, SUSHIL G.  
;; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO  
;; NUMBER OF SEQUENCES: 89  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Abbott Laboratories  
;; STREET: 100 Abbott Park Road  
;; CITY: Abbott Park  
;; STATE: IL  
;; COUNTRY: USA  
;; ZIP: 60064-3500  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: MS-DOS (Windows 95)  
;; SOFTWARE: Microsoft Word (ASCII format output)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/912,129A  
;; FILING DATE: 15-AUG-1997  
;; CLASSIFICATION: 436  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Danckers, Andreas M.  
;; REGISTRATION NUMBER: 32,652  
;; REFERENCE/DOCKET NUMBER: 6109.US.01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 847-937-9803  
;; TELEFAX: 847-938-2623  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 58:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 215 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-912-129A-58

Query Match 86.6%; Score 110; DB 2; Length 215;  
Best Local Similarity 77.3%; Pred. No. 1.5e-08;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQVNET 22  
||:||||:||||:||||:||||:  
Db 127 LLNLWGCKRGLICYTSVRWNET 148

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-50

Query Match      86.6%; Score 110; DB 2; Length 490;
Best Local Similarity 77.3%; Pred. No. 3.6e-08;
Matches 17; Conservative 4; Mismatches 1; Indels

QY      1 LLSSWGCRGLVCYTSVKWNET 22
      ||:||||:||||:||||:
DB      372 LLNLWGCKGRGLICYTSVKWNET 393

RESULT 11
US-08-912-129A-54
; Sequence 54, Application US/08912129A
; Patent No. 5922533
; GENERAL INFORMATION:
; APPLICANT: VALLARI, ANADRUZELA S.
; APPLICANT: HACKETT, JOHN JR.
; APPLICANT: HICKMAN, ROBERT K.
; APPLICANT: VARITEK, VINCENT A. JR.
; APPLICANT: NECKLAWS, ELIZABETH A.
; APPLICANT: GOLDEN, ALAN M.
; APPLICANT: BRENNAN, CATHERINE A.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS (Windows 95)
; SOFTWARE: Microsoft Word (ASCII format output)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,129A
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dancigers, Andreas M.
; REGISTRATION NUMBER: 32,652
; REFERENCE/DOCKET NUMBER: 6109.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-937-9803
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-912-129A-54

```

```

/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dancakers, Andreas M.
/ REGISTRATION NUMBER: 32,652
/ REFERENCE/DOCKET NUMBER: 6109.US.01
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 847-937-9803
/ TELEFAX: 847-938-2623
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 61:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 873 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Protein
/ US-08-912-129A-61

Query Match 86.6%; Score 110; DB 2; Length 873;
Best Local Similarity 77.3%; Pred. No. 6.5e-08;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWCGRLVCYTSVQWNET 22
   ||:||||:||||:||||
Db 601 LLNLWGCKGRLCYTSVKWNET 622

RESULT 14
US-08-817-441-94
; Sequence 94, Application US/08817441
; Patent No. 639294
; GENERAL INFORMATION:
; APPLICANT: CHARNEAU, PIERRE
; APPLICANT: CLAVEL, FRANCOISE
; APPLICANT: BORMAN, ANDREW
; APPLICANT: QUILLIENT, CAROLINE
; APPLICANT: GUETARD, DENISE
; APPLICANT: MONTAGNIER, LUC
; APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE
; APPLICANT: COHEN, JACQUES
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
; TITLE OF INVENTION: SUBTYPE) ANTIGENS
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,441
; FILING DATE: 11-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 95/01391
; FILING DATE: 20-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9412554
; FILING DATE: 20-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9502526
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03260.5005-00000

```

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 94:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-817-441-94

Query Match 85.8%; Score 109; DB 4; Length 37;  
Best Local Similarity 77.3%; Pred. No. 3.5e-09;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22  
||:||||:|||||||  
Db 14 LLNLWGCGRKAICYTSVQWNET 35

## RESULT 15

US-08-817-441-100  
Sequence 100, Application US/08817441  
Patent No. 6399294

## GENERAL INFORMATION:

APPLICANT: CHARNEAU, PIERRE  
APPLICANT: CLAVEL, FRANCOISE  
APPLICANT: BORMAN, ANDREW  
APPLICANT: QUILLENT, CAROLINE  
APPLICANT: GUETARD, DENISE  
APPLICANT: MONTAGNIER, LUC  
APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE  
APPLICANT: COHEN, JACQUES  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
TITLE OF INVENTION: SUBTYPE) ANTIGENS  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA

ZIP: 20005-3315

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,441  
FILING DATE: 11-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR 95/01391  
FILING DATE: 20-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9412554  
FILING DATE: 20-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9502526  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-817-441-100

Query Match 85.8%; Score 109; DB 4; Length 104;  
Best Local Similarity 77.3%; Pred. No. 1e-08;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22  
||:||||:|||||||  
Db 45 LLNLWGCGRKAICYTSVQWNET 66

Search completed: May 7, 2004, 17:53:26  
Job time : 13.1092 secs

61000 11/10/10 11/10/10

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 30.7815 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-8

Perfect score: 127

Sequence: 1 LLSSWCGRGLVCYTSVQWNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	86.6	113	14	US-10-320-786-40
2	110	86.6	215	8	Sequence 40, Appl
3	110	86.6	245	8	Sequence 58, Appl
4	110	86.6	281	8	Sequence 48, Appl
5	110	86.6	373	8	Sequence 120, Appl
6	110	86.6	460	8	Sequence 52, Appl
7	110	86.6	488	8	Sequence 60, Appl
8	110	86.6	490	8	Sequence 95, Appl
9	110	86.6	526	8	Sequence 50, Appl
10	110	86.6	618	8	Sequence 97, Appl
11	110	86.6	706	8	Sequence 54, Appl
12	110	86.6	715	14	Sequence 93, Appl
13	110	86.6	736	8	Sequence 134, Appl
14	110	86.6	873	8	Sequence 91, Appl
15	109	85.8	37	14	Sequence 61, Appl
					Sequence 94, Appl

16	109	85.8	104	14	US-10-026-741-100	Sequence 100, Appl
17	107	84.3	113	14	US-10-320-786-28	Sequence 28, Appl
18	107	84.3	116	14	US-10-320-786-20	Sequence 20, Appl
19	107	84.3	356	14	US-10-357-400-12	Sequence 12, Appl
20	106	83.5	37	14	US-10-026-741-86	Sequence 86, Appl
21	106	83.5	113	14	US-10-320-786-2	Sequence 2, Appl
22	106	83.5	113	14	US-10-320-786-18	Sequence 18, Appl
23	106	83.5	117	14	US-10-320-786-6	Sequence 6, Appl
24	106	83.5	200	9	US-09-854-816-104	Sequence 104, Appl
25	106	83.5	862	14	US-10-369-294-15	Sequence 15, Appl
26	105	82.7	23	14	US-10-026-741-30	Sequence 30, Appl
27	105	82.7	23	14	US-10-320-786-137	Sequence 137, Appl
28	105	82.7	110	14	US-10-320-786-14	Sequence 14, Appl
29	105	82.7	115	14	US-10-320-786-24	Sequence 24, Appl
30	105	82.7	116	14	US-10-320-786-22	Sequence 22, Appl
31	104	81.9	23	14	US-10-320-786-95	Sequence 95, Appl
32	104	81.9	110	14	US-10-320-786-16	Sequence 16, Appl
33	104	81.9	110	14	US-10-320-786-30	Sequence 30, Appl
34	104	81.9	113	14	US-10-320-786-10	Sequence 10, Appl
35	104	81.9	113	14	US-10-320-786-12	Sequence 12, Appl
36	102	80.3	24	14	US-10-320-786-92	Sequence 92, Appl
37	101	79.5	24	14	US-10-320-786-138	Sequence 138, Appl
38	100	78.7	23	14	US-10-320-786-102	Sequence 102, Appl
39	100	78.7	35	14	US-10-026-741-101	Sequence 101, Appl
40	100	78.7	37	14	US-10-026-741-7	Sequence 7, Appl
41	100	78.7	37	14	US-10-026-741-90	Sequence 90, Appl
42	100	78.7	110	14	US-10-320-786-38	Sequence 38, Appl
43	100	78.7	351	14	US-10-026-741-47	Sequence 47, Appl
44	100	78.7	877	14	US-10-026-741-102	Sequence 102, Appl
45	99	78.0	23	14	US-10-320-786-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1  
US-10-320-786-40  
; Sequence 40, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Feeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 40  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Misc\_feature  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: Xaa = unknown  
US-10-320-786-40

Query Match 86.6%; Score 110; DB 14; Length 113;  
Best Local Similarity 77.3%; Pred. No. 1.8e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWCGRGLVCYTSVQWNET 22

Db 39 LLNLWCKGRLLCYTSVQWNET 60



```

; Sequence 60, Application US/08911824
; Publication No. US2003004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN
; TITLE OF INVENTION: DETECTION AND
; FILE REFERENCE: 6165 US 01
; CURRENT APPLICATION NUMBER: US/08/9
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Vers
; SEQ ID NO 60
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Human Immunodeficiency V
; FEATURE:
; OTHER INFORMATION: Encodes recombi
US-08-911-824-60

```

Query Match 86.6%; Score 110; DB 8; Length 460;  
Best Local Similarity 77.3%; Pred. No. 6.7e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels

```
Qy      1 LLSSWGCRGLVCYTSVQWNET 22  
        ||:||:||:||:||:||:  
Db     372 LNLWGCKGRLICYSVKWNET 393
```

## RESULT 7

```

US-08-911-824-95
: Sequence 95, Application US/08911824
: Publication No. US20030004323A1
: GENERAL INFORMATION:
: APPLICANT: Abbott Laboratories
: APPLICANT: Hackett, John R., Jr.
: APPLICANT: Yamaguchi, Julie
: APPLICANT: Golden, Alan M.
: APPLICANT: Brennan, Catherine A.
: APPLICANT: Hickman, Robert K.
: APPLICANT: Devare, Sushil G.
: TITLE OF INVENTION: DETECTION ANTIGEN
: TITLE OF INVENTION: NOVEL ANTIGEN
: FILE REFERENCE: 6165.US.01
: CURRENT APPLICATION NUMBER: US/08/9
: CURRENT FILING DATE: 1997-08-15
: NUMBER OF SEQ ID NOS: 121
: SOFTWARE: FastSeq for Windows Vers
: SEQ ID NO 95
: LENGTH: 488
: TYPE: PRT
: ORGANISM: Human Immunodeficiency V
: FEATURE:
: OTHER INFORMATION: Encodes recombi
US-08-911-824-95

```

Query Match 86.6%; Score 110; DB 8; Length 488;  
Best Local Similarity 77.3%; Pred. No. 7e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels

```
Qy      1 LLSSWGCRGRLVCYTSVQWNET 22
        ||:||:||:||:||:||:||
Db     127 LNLWGCKGRLCYTSVKWNET 148
```

## RESULT 8

US-08-911-824-50  
; Sequence 50, Application US/08911824

```

; Publication No. US20030004232A1
;
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGENIC
; TITLE OF INVENTION: DETECTION AND
; FILE REFERENCE: 6165.US.O1
; CURRENT APPLICATION NUMBER: US/08
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Ver
;
; SEQ ID NO 50
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human Immunodeficiency
; FEATURE:
; OTHER INFORMATION: Encodes recom
; US-08-911-824-50

```

Query Match 86.6%; Score 110; DB 8; Length 490;  
Best Local Similarity 77.3%; Pred. No. 7.1e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels

```
Qy      1 LLSSWGCRGLVCYTSVQWNET 22  
        ||:||:||:||:||:||:  
Db     372 LNLWGCKGRLLCYTSVKWNET 393
```

## RESULT 9

```

US-08-911-824-97
; Sequence 97, Application US/08911824
; Publication No. US20030004323A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Hackett, John R., Jr.
; APPLICANT: Yamaguchi, Julie
; APPLICANT: Golden, Alan M.
; APPLICANT: Brennan, Catherine A.
; APPLICANT: Hickman, Robert K.
; APPLICANT: Devare, Sushil G.
; TITLE OF INVENTION: NOVEL ANTIGEN
; TITLE OF INVENTION: DETECTION AND
; FILE REFERENCE: 6165_US.O1
; CURRENT APPLICATION NUMBER: US/08/9
; CURRENT FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Vers
; SEQ ID NO 97
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Human Immunodeficiency V
; FEATURE:
; OTHER INFORMATION: Encodes recombi
US-08-911-824-97

```

Query Match 86.6%; Score 110; DB 8; Length 526;  
Best Local Similarity 77.3%; Pred. No. 7.5e-07;  
Matches 17; Conservative 4; Mismatches 1; Indels

```

QY      1  LLSSWGCRGLVCYTSVQWNET  22
        ||: |||: |||: |||: |||: |||
Db     372  LLNLWGCKGRGLCYTSVQWNET  393

```

RESULT 10

US-08-911-824-54  
; Sequence 54, Application US/08911824  
; Publication No. US20030004323A1

Query Match	Best Local Similarity	Score	DB	Length	Indels	Mismatches	Gaps
Query Match	86.6%;	Score 110;	DB 8;	Length 618;			
Best Local Similarity	77.3%;	Pred. No. 8.7e-07;					
Matches	17;	Conservative	4;	Mismatches	1;	Indels	0;
Qy	1	LLSSWGCGRGLVCYTSVQWNET 22					
Db	372	LLNLWGCKGRGLCYTSVKWNET 393					
RESULT 11							
US-08-911-824-93							
Sequence 93, Application US/08911824							
Publication No. US20030004323A1							
GENERAL INFORMATION:							
APPLICANT: Abbott Laboratories							
APPLICANT: Hackett, John R., Jr.							
APPLICANT: Yamaguchi, Julie							
APPLICANT: Golden, Alan M.							
APPLICANT: Brennan, Catherine A.							
APPLICANT: Hickman, Robert K.							
APPLICANT: Devare, Sushil G.							
TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE							
DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV							
FILE REFERENCE: 6165.US.01							
CURRENT APPLICATION NUMBER: US/08/911,824							
CURRENT FILING DATE: 1997-08-15							
NUMBER OF SEQ ID NOS: 121							
SOFTWARE: FastSeq for Windows Version 3.0							
SEQ ID NO 93							
LENGTH: 706							
TYPE: PRT							
ORGANISM: Human Immunodeficiency Virus							
FEATURE:							
OTHER INFORMATION: Encodes recombinant protein pCO-13CKS							
US-08-911-824-93							
Query Match	86.6%;	Score 110;	DB 8;	Length 706;			
Best Local Similarity	77.3%;	Pred. No. 9.9e-07;					
Matches	17;	Conservative	4;	Mismatches	1;	Indels	0;
Qy	1	LLSSWGCGRGLVCYTSVQWNET 22					
Db	618	LLNLWGCKGRGLCYTSVKWNET 639					
RESULT 12							
US-10-320-786-134							
Sequence 134, Application US/10320786							
Publication No. US20030180759A1							
GENERAL INFORMATION:							
APPLICANT: Abbott Laboratories							
APPLICANT: Hackett, John R., Jr.							
APPLICANT: Yamaguchi, Julie							
APPLICANT: Golden, Alan M.							
APPLICANT: Brennan, Catherine A.							
APPLICANT: Hickman, Robert K.							
APPLICANT: Devare, Sushil G.							
TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE							
DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV							
FILE REFERENCE: 6165.US.01							
CURRENT APPLICATION NUMBER: US/08/911,824							
CURRENT FILING DATE: 1997-08-15							
NUMBER OF SEQ ID NOS: 121							
SOFTWARE: FastSeq for Windows Version 3.0							
SEQ ID NO 93							
LENGTH: 706							
TYPE: PRT							
ORGANISM: Human Immunodeficiency Virus							
FEATURE:							
OTHER INFORMATION: Encodes recombinant protein pCO-13CKS							
US-08-911-824-93							
Query Match	86.6%;	Score 110;	DB 8;	Length 736;			
Best Local Similarity	77.3%;	Pred. No. 1e-06;					
Matches	17;	Conservative	4;	Mismatches	1;	Indels	0;
Qy	1	LLSSWGCGRGLVCYTSVQWNET 22					



APPLICANT: Golden, Alan M.  
APPLICANT: Brennan, Catherine A.  
APPLICANT: Hickman, Robert K.  
APPLICANT: Devare, Sushil G.  
TITLE OF INVENTION: NOVEL ANTIGEN CONSTRUCTS USEFUL IN THE  
TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF ANTIBODIES TO HIV  
FILE REFERENCE: 6165.US.01  
CURRENT APPLICATION NUMBER: US/08/911,824  
CURRENT FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 61  
LENGTH: 873  
TYPE: PRT  
ORGANISM: Human Immunodeficiency Virus  
FEATURE:  
OTHER INFORMATION: HIV-1 Group O isolate HAM112  
US-08-911-824-61  
Query Match 86.8%; Score 110; DB 8; Length 873;  
Best Local Similarity 77.3%; Pred. No. 1.2e-06;  
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 1 LLSSWGCGRGLVCYTSVQWNET 22  
Db 601 LLNLWGCKRGLCYTSVKWNET 622  
RESULT 15  
US-10-026-741-94  
Sequence 94, Application US/10026741  
Publication No. US20030049604A1  
GENERAL INFORMATION:  
APPLICANT: CHARNEAU, PIERRE  
BORMAN, ANDREW  
CLAVEL, FRANCOISE  
QUILLET, CAROLINE  
GUETARD, DENISE  
MONTAGNIER, LUC  
DONJON DE SAINT-MARTIN, JACQUELINE  
COHEN, JACQUES  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
SUBTYPE) ANTIGENS  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/026,741  
FILING DATE: 27-Dec-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/817,441  
FILING DATE: 31-AUG-1998  
APPLICATION NUMBER: PCT/FR 95/01391  
FILING DATE: 20-OCT-1995  
APPLICATION NUMBER: FR 9412554  
FILING DATE: 20-OCT-1994  
APPLICATION NUMBER: FR 9502526  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 94:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 94:  
US-10-026-741-94  
Query Match 85.8%; Score 109; DB 14; Length 37;  
Best Local Similarity 77.3%; Pred. No. 8.9e-08;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 LLSSWGCGRGLVCYTSVQWNET 22  
Db 14 LLNLWGCKRGLCYTSVQWNET 35  
Search completed: May 7, 2004, 18:29:21  
Job time : 31.7815 secs

6400 11/12/19 11/12/19 11/12/19

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 9.2437 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-8  
Perfect score: 127  
Sequence: 1 LLSSWGCGRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	85.8	104	2 S52930	GP41 ENV protein -
2	106	83.5	863	2 A53034	gag polyprotein -
3	100	78.7	877	2 S49197	envelope protein p
4	83	65.4	854	1 VCLJSI	env polyprotein pr
5	82	64.6	357	2 S21990	envelope protein g
6	82	64.6	358	2 S22002	envelope protein g
7	81	63.8	358	2 S22000	envelope protein g
8	81	63.8	358	2 S70417	envelope protein g
9	80	63.0	443	2 A41621	env polyprotein P
10	80	63.0	833	2 S54384	envelope polyprote
11	80	63.0	855	1 VCLJZR	env polyprotein pr
12	79	62.2	357	2 S22006	envelope protein g
13	79	62.2	357	2 S21994	envelope protein g
14	79	62.2	357	2 S22004	envelope protein g
15	79	62.2	357	2 S21996	envelope protein g
16	79	62.2	357	2 S21992	envelope protein g
17	79	62.2	358	2 S21998	envelope protein g
18	79	62.2	445	2 A41621	env polyprotein D
19	79	62.2	454	2 B41621	env polyprotein M
20	79	62.2	843	1 H44001	env polyprotein pr
21	79	62.2	847	2 T09448	envelope glycoprot
22	79	62.2	847	2 S13289	env protein - huma
23	79	62.2	852	1 VCLJBR	env polyprotein -
24	79	62.2	852	2 T12016	envelope glycoprot
25	79	62.2	854	2 S13288	env protein - huma
26	79	62.2	855	1 VCLJAZ	env polyprotein pr
27	79	62.2	856	1 VCLJH3	env polyprotein pr
28	79	62.2	856	1 VCLJVL	env polyprotein pr
29	79	62.2	856	1 VCLJ3W	env polyprotein pr

env polyprotein pr  
env polyprotein pr  
env polyprotein pr  
env polyprotein -  
Env transmembrane  
env polyprotein -  
env polyprotein pr  
env polyprotein pr  
env polyprotein pr  
env protein - huma  
env polyprotein -  
env protein - simi  
env polyprotein pr  
env polyprotein pr  
env protein - huma  
env protein - huma  
env protein - huma

#### ALIGNMENTS

##### RESULT 1

S52930

GP41 ENV protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 06-Jun-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999

C;Accession: S52930

R;Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.;

submitted to the EMBL Data Library, January 1995

A;Description: A novel HIV-1 O strain illustrates the diversity of the O group.

A;Reference number: S52929

A;Accession: S52930

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-104 <COH>

A;Cross-references: EMBL:X84328; NID:g695526; PIDN:CAA59066.1; PID:g695527

C;Superfamily: type E retrovirus env polyprotein

Query Match 85.8%; Score 109; DB 2; Length 104;  
Best Local Similarity 77.3%; Pred. No. 1.1e-08;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22

DB 45 LLNLWGCRGKLCYTSVQWNET 66

##### RESULT 2

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C;Accession: A53034

R;Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborght, B.; van der Groen,

J. Virol. 68, 1586-1596, 1994

A;Title: Genomic cloning and complete sequence analysis of a highly divergent African h

A;Reference number: A53034; MUID:94149849; PMID:8107220

A;Accession: A53034

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-863 <VAN>

A;Cross-references: GB:I02587

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: polyprotein

Query Match 83.5%; Score 106; DB 2; Length 863;  
Best Local Similarity 77.3%; Pred. No. 1.9e-07;  
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22

DB 594 LLNLWGCRGKLCYTSVQWNET 615



```

F:424-443/Domain: transmembrane #status predicted <TM>:
F:9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site:

Query Match      63.0%; Score 80; DB 2; Length 443;
Best Local Similarity 54.5%; Pred. No. 0.00058;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LLSWGCGRGLVCVTSVQWNET 22
Db 332 LLGIWGCGLICTTAVPNSS 353

RESULT 10
S54384
env polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C:Accession: S54384
R:Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54384
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-853 <THE>
A:Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: polyprotein

Query Match      63.0%; Score 80; DB 2; Length 853;
Best Local Similarity 54.5%; Pred. No. 0.001;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LLSWGCGRGLVCVTSVQWNET 22
Db 589 LLGIWGCGLICTTAVPNSS 610

RESULT 11
VCLJZR
env polyprotein precursor - human immunodeficiency virus Zr-6
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus Zr-6
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.;
Gene 52, 71-82, 1987
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot
A:Reference number: A26192; MUID:87248097; PMID:3036660
A:Accession: D26192
A:Molecule type: DNA
A:Residues: 1-855 <SPI>
A:Cross-references: GB:X03458; GB:M16322; NID:g329398; PIDN:AAA45380.1; PID:g329403
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-855/Product: env polyprotein #status predicted <NAT>
F:250-500/Product: exterior membrane glycoprotein #status predicted <EXT>
F:501-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,145,154,158,166,189,199,236,243,264,278,291,297,333,340,355,386,392,398,40

Query Match      63.0%; Score 80; DB 1; Length 855;
Best Local Similarity 54.5%; Pred. No. 0.001;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LLSWGCGRGLVCVTSVQWNET 22
Db 591 LLGIWGCGLICTTAVPNSS 612

RESULT 12

```

```
S22006
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70420; S22006
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61354; NID:G60190; PIDN:CAA43620.1; PID:G60191
A:Experimental source: patient L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match 62.2%; Score 79; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00067;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCGRGLVCYTSVQWN 20
DB 93 LLGIWCGSGKLICTTTVPWN 112

RESULT 13
S21994
envelope protein gp120/gp41 - human immunodeficiency virus type 1
A:Variety: isolate 27B
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21994; S70421
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
A:Accession: S21994
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61355; NID:G60179; PIDN:CAA43622.1; PID:G60180
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70421
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140, 'X', 142-312, 'X', 314-357 <ST2>
A:Cross-references: EMBL:X61355; NID:G60179
C:Superfamily: type E retrovirus env polyprotein

Query Match 62.2%; Score 79; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00067;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCGRGLVCYTSVQWN 20
DB 93 LLGIWCGSGKLICTTTVPWN 112

RESULT 14
S22004
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 4B
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22004; S70419
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S21990
```

```
A:Accession: S22004
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61353; NID:G60188; PIDN:CAA43618.1; PID:G60189
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292, 'X', 294-357 <ST2>
A:Cross-references: EMBL:X61353; NID:G60188
C:Superfamily: type E retrovirus env polyprotein

Query Match 62.2%; Score 79; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00067;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCGRGLVCYTSVQWN 20
DB 93 LLGIWCGSGKLICTTTVPWN 112

RESULT 15
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70422; S21996
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined by monoclonal antibodies
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61356; NID:G60181; PIDN:CAA43624.1; PID:G1067129
A:Experimental source: patient 27L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match 62.2%; Score 79; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 0.00067;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCGRGLVCYTSVQWN 20
DB 93 LLGIWCGSGKLICTTTVPWN 112

Search completed: May 7, 2004, 17:51:02
Job time : 9.2437 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 5.54622 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-8  
Perfect score: 127  
Sequence: 1 LSSWCGRGLVCYTSQVQNET 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	65.4	854	1 ENV_SIVCZ	P17281 chimpanzee
2	80	63.0	853	1 ENV_HV1A2	P12487 human immun
3	80	63.0	855	1 ENV_HV1A6	P04580 human immun
4	79	62.2	843	1 ENV_HV1Y2	P35961 human immun
5	79	62.2	847	1 ENV_HV1S1	P19550 human immun
6	79	62.2	847	1 ENV_HV1W2	P05880 human immun
7	79	62.2	851	1 ENV_HV1B8	P04582 human immun
8	79	62.2	852	1 ENV_HV1B9	P12488 human immun
9	79	62.2	852	1 ENV_HV1S3	P19549 human immun
10	79	62.2	853	1 ENV_HV1MF	P19551 human immun
11	79	62.2	855	1 ENV_HV1A2	P03378 human immun
12	79	62.2	855	1 ENV_HV1OY	P20888 human immun
13	79	62.2	856	1 ENV_HV1B1	P03375 human immun
14	79	62.2	856	1 ENV_HV1H2	P04578 human immun
15	79	62.2	856	1 ENV_HV1LW	Q70626 human immun
16	79	62.2	856	1 ENV_HV1MN	P05877 human immun
17	79	62.2	856	1 ENV_HV1PV	P03376 human immun
18	79	62.2	856	1 ENV_HV1SC	P05878 human immun
19	79	62.2	856	1 ENV_HV1W1	P31872 human immun
20	79	62.2	861	1 ENV_HV1BR	P03377 human immun
21	79	62.2	865	1 ENV_HV1RH	P04579 human immun
22	79	62.2	867	1 ENV_HV1J3	P12489 human immun
23	79	62.2	868	1 ENV_HV1C4	P05879 human immun
24	78	61.4	821	1 ENV_SIVGB	P22380 simian immu
25	77	60.6	848	1 ENV_HV1JR	P20871 human immun
26	77	60.6	856	1 ENV_HV1H3	P04624 human immun
27	77	60.6	856	1 ENV_HV1ZH	P05881 human immun
28	76	59.8	846	1 ENV_HV1ND	P18799 human immun
29	74.5	58.7	856	1 ENV_HV2N2	P05883 human immun
30	74.5	58.7	859	1 ENV_HV2D2	P15831 human immun
31	74.5	58.7	885	1 ENV_SIVS4	P12492 simian immu
32	73	57.5	853	1 ENV_HV1EL	P04581 human immun
33	73	57.5	857	1 ENV_HV2KR	Q74126 human immun

## ALIGNMENTS

### RESULT 1

ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
AC	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retrovirdae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1.";			
RL	Nature 345:356-359(1990).			
CC	-!- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).			
DR	EMBL; X52154; CAA36407.1; -			
DR	PIR; S09990; VCLJ31.			
DR	HIV; X52154; ENVSCPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .) (POTENTIAL).

34	73	57.5	861	1 ENV_HV1KB	P31819 human immun
35	73	57.5	863	1 ENV_HV1Z8	P05882 human immun
36	71.5	56.3	712	1 ENV_HV2S2	P32536 human immun
37	71.5	56.3	846	1 ENV_HV2SB	P12449 human immun
38	71.5	56.3	859	1 ENV_HV2ST	P20872 human immun
39	71	55.9	859	1 ENV_HV1MA	P04583 human immun
40	70.5	55.5	859	1 ENV_HV2CA	P24105 human immun
41	70	55.1	865	1 ENV_SIVAT	P05886 simian immu
42	69	54.3	851	1 ENV_HV2D1	P17755 human immun
43	69	54.3	851	1 ENV_HV2G1	P18040 human immun
44	69	54.3	858	1 ENV_HV2RO	P04577 human immun
45	69	54.3	860	1 ENV_HV2BE	P18094 human immun

```
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 65.4%; Score 83; DB 1; Length 854;
Best Local Similarity 54.5%; Pred. No. 4.1e-05;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQVNET 22
DB 582 ILGLWGCSGKAVCYTTPWNSS 603

RESULT 2
ENV_HV122
ID ENV_HV122 STANDARD; PRT; 853 AA.
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV5226.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 833 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```
FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCAFF7008 CRC64;

Query Match 63.0%; Score 80; DB 1; Length 853;
Best Local Similarity 54.5%; Pred. No. 0.00012;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQVNET 22
DB 589 LLGIWGCSGKLCITTPWNSS 610

RESULT 3
ENV_HV126
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Peorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
CC Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.
```



DR PIR; D26192; VCLJZR.  
 DR HIV; K03458; ENV526.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 207 BY SIMILARITY.  
 FT DISULFID 125 198 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 220 249 BY SIMILARITY.  
 FT DISULFID 230 241 BY SIMILARITY.  
 FT DISULFID 298 332 BY SIMILARITY.  
 FT DISULFID 378 444 BY SIMILARITY.  
 FT DISULFID 385 417 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;  
 Query Match 63.0%; Score 80; DB 1; Length 855;  
 Best Local Similarity 54.5%; Pred. No. 0.00012;  
 Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 LLSSWCGRGLVCYTSQVNNET 22  
 DB 591 LUGIWGSGKLICTTTPMNS 612  
 RESULT 4  
 ID ENV\_HV1Y2 STANDARD; PRT; 843 AA.  
 AC F35961;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.

OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36377;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93021387; PubMed=1404605;  
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,  
 Shaw G.M.;  
 RA "Complete nucleotide sequence, genome organization, and biological  
 RT properties of human immunodeficiency virus type 1 in vivo: evidence  
 RT for limited defectiveness and complementation.";  
 RL J. Virol. 66:6587-6600(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M93258; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; H44001; H44001.  
 DR PDB; 1G9N; 27-DEC-00.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.  
 FT TRANSMEM POTENTIAL.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 201 BY SIMILARITY.  
 FT DISULFID 125 192 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 214 243 BY SIMILARITY.  
 FT DISULFID 224 235 BY SIMILARITY.  
 FT DISULFID 292 326 BY SIMILARITY.  
 FT DISULFID 373 432 BY SIMILARITY.  
 FT DISULFID 380 405 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

```

Query Match          62.2%; Score 79; DB 1; Length 843;
Best Local Similarity 60.0%; Pred. No. 0.00017;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQWN 20
    |||||:|:|:|:|
Db 579 LLGIWCGSKLICTTAVPNW 598

RESULT 5
ENV_HV1S1
ID - ENV HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65024; AAA45072.1; -.
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENV$SF162.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match          62.2%; Score 79; DB 1; Length 847;
Best Local Similarity 60.0%; Pred. No. 0.00017;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWGCRGLVCYTSVQWN 20
    |||||:|:|:|:|
Db 583 LLGIWCGSKLICTTAVPNW 602

RESULT 6
ENV_HV1W2
ID - ENV HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Skaal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT at risk for AIDS.";
RL Science 232:1548-1553(1986).
CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12507; AAB12990.1; -.
DR HIV; M12507; ENV$WMJ2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 202 BY SIMILARITY.
FT DISULFID 125 193 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 215 244 BY SIMILARITY.
FT DISULFID 225 236 BY SIMILARITY.
FT DISULFID 293 326 BY SIMILARITY.

```





```
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 96663 MW; EE7BBF8D23C9910D CRC64;

Query Match 62.2%; Score 79; DB 1; Length 852;
Best Local Similarity 60.0%; Pred. No. 0.00017;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWCGRGLVCYTSVQWN 20
Db 588 LLGIMGCGKLICTTAVPW 607

RESULT 10
ENV_HV1MF
ID ENV_HV1MF STANDARD; PRT; 853 AA.
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonica C., Mann A.M., Meier C.,
Wasiak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
variants diminished in the ability to induce syncytium-independent
cytolysis."
RL J. Virol. 64:3792-3803(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M33943; AAA44850.1; --
CC PDB; 1AIK; 16-JUN-97.
CC HIV; M33943; ENVSMFA.
CC InterPro; IPR000328; Env GP41.
CC Dr InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 30
```

```
FT CHAIN 31 509
FT CHAIN 510 853
FT DISULFID 54 74
FT DISULFID 119 203
FT DISULFID 126 194
FT DISULFID 131 157
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 329
FT DISULFID 376 443
FT DISULFID 383 416
FT CARBOHYD 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 195 195
FT CARBOHYD 232 232
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 287 287
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 330 330
FT CARBOHYD 354 354
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 395 395
FT CARBOHYD 404 404
FT CARBOHYD 446 446
FT CARBOHYD 461 461
FT CARBOHYD 609 609
FT CARBOHYD 614 614
FT CARBOHYD 623 623
FT CARBOHYD 635 635
FT CARBOHYD 672 672
FT CARBOHYD 748 748
FT CARBOHYD 814 814
SQ SEQUENCE 853 AA; 96912 MW; 3377B993B6F22ABA CRC64;

Query Match 62.2%; Score 79; DB 1; Length 853;
Best Local Similarity 60.0%; Pred. No. 0.00017;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSSWCGRGLVCYTSVQWN 20
Db 590 LLGIMGCGKLICTTAVPW 609

RESULT 11
ENV_HV1A2
ID ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
(RARV-2).";
RL Science 227:484-492(1985).
```

----- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)). -----

EMBL; K02007; AAB59882.1; --  
 DR PIR; A03976; VCLJ2A2.  
 DR HIV; K02007; ENV5SF2.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 208 BY SIMILARITY.  
 FT DISULFID 125 199 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 211 250 BY SIMILARITY.  
 FT DISULFID 231 242 BY SIMILARITY.  
 FT DISULFID 239 333 BY SIMILARITY.  
 FT DISULFID 380 442 BY SIMILARITY.  
 FT DISULFID 387 415 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 815 815 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 62.2%; Score 79; DB 1; Length 855;  
 Best Local Similarity 60.0%; Pred. No. 0.00017;  
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWN 20

Db 591 LLGIWGCGKLCITTAVPWN 610

RESULT 12

ENV HV10Y STANDARD; PRT; 855 AA.  
 ID ENV HV10Y  
 AC P20888;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV  
 OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90148544; PubMed=2559749;  
 RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;  
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese individual presenting an atypical western blot.";  
 RL AIDS 3:707-715 (1989).  
 CC -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A HEALTHY GABONESE INDIVIDUAL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)). -----

EMBL; M26727; AAA83397.1; --  
 DR HIV; M26727; ENV5OVI.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 509 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 510 855 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 210 BY SIMILARITY.  
 FT DISULFID 125 201 BY SIMILARITY.  
 FT DISULFID 130 162 BY SIMILARITY.  
 FT DISULFID 223 252 BY SIMILARITY.  
 FT DISULFID 233 244 BY SIMILARITY.  
 FT DISULFID 301 335 BY SIMILARITY.  
 FT DISULFID 381 442 BY SIMILARITY.  
 FT DISULFID 388 415 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).











[illegible]

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 27.916 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-8

Perfect score: 127  
Sequence: 1 LLSSWGCGRGLVCYTSVQWNET 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	92.1	216	15 Q9IEC5	Q9iec5 human immun
2	114	89.8	116	15 Q7ZJN9	Q7zjn9 human immun
3	114	89.8	135	15 Q9DQL9	Q9dql9 human immun
4	113	89.0	209	15 Q9IEB6	Q9ieb6 human immun
5	112	88.2	219	15 Q9IEB6	Q9ieb6 human immun
6	110	86.6	116	15 O40458	O40458 human immun
7	110	86.6	124	15 Q9IHU7	Q9ihu7 human immun
8	110	86.6	126	15 Q9IHV1	Q9ihv1 human immun
9	110	86.6	130	15 Q9IHU9	Q9ihu9 human immun
10	110	86.6	158	15 Q8J3N6	Q8j3n6 human immun
11	110	86.6	218	15 Q9IE95	Q9ie95 human immun
12	110	86.6	230	15 Q9IEB2	Q9ieb2 human immun
13	110	86.6	240	15 Q9IE32	Q9ie32 human immun
14	110	86.6	536	15 Q9IEE5	Q9iee5 human immun
15	110	86.6	538	15 Q9IED5	Q9ied5 human immun
16	110	86.6	872	15 Q8Q7H0	Q8q7h0 human immun

17	110	86.6	879	15 Q9WIU9	Q9wiu9 human immun
18	110	86.6	880	15 Q8Q7H1	Q8q7h1 human immun
19	110	86.6	882	15 Q8Q7F9	Q8q7f9 human immun
20	110	86.6	887	15 Q8Q7H6	Q8q7h6 human immun
21	110	86.6	887	15 Q8Q7G9	Q8q7g9 human immun
22	110	86.6	900	15 Q9QNZ8	Q9qnz8 human immun
23	109	85.8	104	15 Q76163	Q76163 human immun
24	109	85.8	216	15 Q9IEC7	Q9iec7 human immun
25	109	85.8	535	15 Q9IEF2	Q9ief2 human immun
26	108	85.0	125	15 Q9IHU8	Q9ihu8 human immun
27	108	85.0	230	15 Q9IEA9	Q9iea9 human immun
28	108	85.0	238	15 Q9DIK1	Q9dik1 human immun
29	108	85.0	242	15 Q9IE31	Q9ie31 human immun
30	108	85.0	529	15 Q9IEE2	Q9iee2 human immun
31	107	84.3	114	15 O40448	O40448 human immun
32	107	84.3	116	15 O40459	O40459 human immun
33	107	84.3	137	15 Q9IHV5	Q9ihv5 human immun
34	107	84.3	155	15 Q8J3Q4	Q8j3q4 human immun
35	107	84.3	172	15 Q9IEB3	Q9ieb3 human immun
36	107	84.3	177	15 Q9IEB0	Q9ieb0 human immun
37	107	84.3	208	15 Q9IEA3	Q9iea3 human immun
38	107	84.3	219	15 Q9IEC8	Q9iec8 human immun
39	107	84.3	227	15 Q9IE99	Q9ie99 human immun
40	107	84.3	234	15 Q9IEC2	Q9iec2 human immun
41	107	84.3	342	15 O11942	O11942 human immun
42	107	84.3	418	15 O36547	O36547 human immun
43	107	84.3	502	15 Q9IEE3	Q9iee3 human immun
44	107	84.3	871	15 O57074	O57074 human immun
45	107	84.3	871	15 Q8Q7I2	Q8q7i2 human immun

ALIGNMENTS

RESULT 1

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5  
AC Q9IEC5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Gp41 (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF07;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RA Philippe M.;  
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ236394; CAB96243.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
DR Transmembrane.  
KW NON\_TER  
FT NON\_TER 1 1  
SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 92.1%; Score 117; DB 15; Length 216;

Best Local Similarity 86.4%; Pred. No. 2.6e-10;

Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22

Db 47 LLSSWGCGRGLVCYTSVQWNET 68

```
RESULT 2
Q7ZJN9 ID Q7ZJN9 PRELIMINARY; PRT; 116 AA.
AC Q7ZJN9;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08692A;
RA Swanson P.A., Devare S.G., Hackett J.R. Jr.;
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M
RT (Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol
RT integrase, and env gp41.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF214120; A061840.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 116
FT NON_TER 116
SQ SEQUENCE 116 AA; 13865 MW; AF1DFC57E059061 CRC64;

Query Match 89.8%; Score 114; DB 15; Length 116;
Best Local Similarity 81.8%; Pred. No. 4.2e-10;
Matches 18; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22
DB 32 LLSWGCGRLVCYTSVKWNQT 53

RESULT 3
Q9DQL9 ID Q9DQL9 PRELIMINARY; PRT; 135 AA.
AC Q9DQL9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-M1321;
RX MEDLINE=20584646; PubMed=11153079;
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,
RA Vellido J., Salas A., Moreno A., Garcia-Saiz A.;
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa
RT in Spain.";
RL AIDS Res. Hum. Retroviruses 16:1967-1971 (2000).
DR EMBL; AF255939; AAG36894.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 135
FT NON_TER 135
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF7718BB36 CRC64;

Query Match 89.8%; Score 114; DB 15; Length 135;
Best Local Similarity 81.8%; Pred. No. 4.8e-10;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22
```

---

```
Db 48 LLSWGCGRLVCYTSVKWNST 69

RESULT 4
Q9IEB6 ID Q9IEB6 PRELIMINARY; PRT; 209 AA.
AC Q9IEB6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Env polypeptide, gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF107;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA Deplenne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ238862; CAB96300.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 209
FT NON_TER 209
SQ SEQUENCE 209 AA; 24116 MW; 91AC9BE2BFE74B83 CRC64;

Query Match 89.0%; Score 113; DB 15; Length 209;
Best Local Similarity 81.8%; Pred. No. 1.1e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCGRLVCYTSVQWNET 22
DB 56 LLSWGCGRLVCYTSVKWNST 77

RESULT 5
Q9IEB6 ID Q9IEB6 PRELIMINARY; PRT; 219 AA.
AC Q9IEB6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCF14;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ236404; CAB96252.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 219
FT NON_TER 219
SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;
```

```
Query Match      88.2%; Score 112; DB 15; Length 219;
Best Local Similarity 81.8%; Pred. No. 1.8e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
   ||:||||:||||:||||:||||:
Db 57 LLNSWCKGKGLVCYTSVKWNT 78

RESULT 6
O40458 PRELIMINARY; PRT; 116 AA.
AC O40458;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope transmembrane glycoprotein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=group O;
RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;
RT "Molecular characterization of envelope transmembrane glycoprotein of
RT 14 new human immunodeficiency virus type 1 group O strains from
RT different African countries.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Y09774; CAA70913.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; 540F5B3B7AD849D8 CRC64;

Query Match      86.6%; Score 110; DB 15; Length 116;
Best Local Similarity 77.3%; Pred. No. 1.8e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
   ||:||||:||||:||||:||||:
Db 39 LLNLWGCKGRLVCYTSVQWNT 60

RESULT 7
Q9IHU7 PRELIMINARY; PRT; 124 AA.
AC Q9IHU7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM768;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229237; AAF71914.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13865 MW; 540F5B3B7AD849D8 CRC64;

Query Match      86.6%; Score 110; DB 15; Length 126;
Best Local Similarity 81.8%; Pred. No. 1.9e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
   ||:||||:||||:||||:||||:
Db 40 LLNLWGCKGRLVCYTSVKWNS 61

RESULT 9
Q9IHU9 PRELIMINARY; PRT; 130 AA.
AC Q9IHU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=10933623;
```

```
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 15187 MW; 105D515F114450F8 CRC64;

Query Match      86.6%; Score 110; DB 15; Length 124;
Best Local Similarity 77.3%; Pred. No. 1.9e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
   ||:||||:||||:||||:||||:
Db 33 LLNLWGCKGRLVCYTSVKWNET 54

RESULT 8
Q9IHV1 PRELIMINARY; PRT; 126 AA.
AC Q9IHV1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97ES203;
RX MEDLINE=20386754; PubMed=10933623;
RA Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA Pieniazek D., Schable C., Lal R.B.;
RT "Phylogenetic analysis of protease and transmembrane regions of HIV
RT type 1 group O.";
RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR EMBL; AF229233; AAF71910.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 15169 MW; 13FB101ECDFF0DDD CRC64;

Query Match      86.6%; Score 110; DB 15; Length 126;
Best Local Similarity 81.8%; Pred. No. 1.9e-09;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
   ||:||||:||||:||||:||||:
Db 40 LLNLWGCKGRLVCYTSVKWNS 61

RESULT 9
Q9IHU9 PRELIMINARY; PRT; 130 AA.
AC Q9IHU9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN GP41.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CM798;
RX MEDLINE=20386754; PubMed=10933623;
```



```
ID Q9IE32 PRELIMINARY; PRT; 240 AA.
AC Q9IE32;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TM, GP41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP STRAIN=BCF11;
RC Roques P., Robertson D., Sousquiere S., Damond F., Mauciere P.,
RA Depienne C., Brun-Vezinet F., Dormont D., Simon F.O.;
RT "Phylogenetic analysis and subtyping of 47 HIV-1 group O isolates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243365; CAB96335.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
DR Transmembrane. 1 1
FT NON_TER 240 240
FT NON_TER 240 240
SQ SEQUENCE 240 AA; 27723 MW; 1E206E8D491A4197 CRC64;

Query Match 86.6%; Score 110; DB 15; Length 240;
Best Local Similarity 77.3%; Pred. No. 3.6e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
||:||||:||||:||||:
Db 60 LLNLWGCGRGLVCYTSVQWNT 81

RESULT 14
Q9IEE5 PRELIMINARY; PRT; 536 AA.
AC Q9IEE5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP STRAIN=YBF15;
RC Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA Depienne C., Brun-Vezinet F., Dormont D.;
RT "Hiv-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13062; CAB96223.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 536 536
FT NON_TER 536 536
SQ SEQUENCE 536 AA; 60419 MW; 0F533ABAF08FE552 CRC64;

Query Match 86.6%; Score 110; DB 15; Length 536;
Best Local Similarity 77.3%; Pred. No. 8.1e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LLSSWGCGRGLVCYTSVQWNET 22
||:||||:||||:||||:
Db 361 LLTLWGCGRGLVCYTSVQWNET 382

RESULT 15
Q9IED5 PRELIMINARY; PRT; 538 AA.
AC Q9IED5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YBF51;
RA Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA Depienne C., Brun-Vezinet F., Dormont D.;
RT "Hiv-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13072; CAB96233.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 538 538
FT NON_TER 538 538
SQ SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEF1D CRC64;

Query Match 86.6%; Score 110; DB 15; Length 538;
Best Local Similarity 77.3%; Pred. No. 8.1e-09;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCGRGLVCYTSVQWNET 22
||:||||:||||:||||:
Db 374 LLSSWGCGRGLVCYTSVQWNT 395

Search completed: May 7, 2004, 17:49:12
Job time : 27.916 secs
```

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:31:55 ; Search time 30.4538 Seconds  
(without alignments)  
148.447 Million cell updates/sec

Title: US-09-147-362A-9

Perfect score: 91  
Sequence: 1 LLSWCKGRLVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	16	2 AAW80467	Aaw80467 Peptide d
2	91	100.0	22	2 AAW80461	Aaw80461 Peptide d
3	91	100.0	22	2 AAW80462	Aaw80462 Peptide d
4	88	96.7	16	2 AAW80468	Aaw80468 Peptide d
5	88	96.7	22	2 AAW80472	Aaw80472 Peptide d
6	88	96.7	22	2 AAW80466	Aaw80466 Peptide d
7	88	96.7	28	2 AAW80473	Aaw80473 Peptide d
8	88	96.7	33	3 AAB12231	Aab12231 Partial s
9	88	96.7	33	3 AAB12212	Aab12212 Partial s
10	88	96.7	40	2 AAW07346	Aaw07346 Partial s
11	88	96.7	40	2 AAW07352	Aaw07352 Partial s
12	87	95.6	22	2 AAW80463	Aaw80463 Peptide d
13	86	94.5	33	3 AAB12236	Aab12236 Partial s
14	86	94.5	40	2 AAW07343	Aaw07343 Partial s
15	85	93.4	32	2 AAW80469	Aaw80469 Peptide d
16	85	93.4	33	3 AAB12226	Aab12226 Partial s
17	85	93.4	40	2 AAW07348	Aaw07348 Partial s
18	84	92.3	17	2 AAW80474	Aaw80474 Peptide d
19	82	90.1	22	2 AAW80460	Aaw80460 Peptide d
20	82	90.1	23	2 AAY05623	Aay05623 HIV-1 gro
21	82	90.1	23	3 AAB12263	Aab12263 Partial s
22	82	90.1	23	7 ABB80287	Abb80287 Immunogen
23	82	90.1	24	2 AAY05624	Aay05624 HIV-1 gro
24	82	90.1	24	2 AAY05612	Aay05612 HIV-1 gro
25	82	90.1	32	4 AAG65199	Aag65199 Human imm

NP-CPX

ALIGNMENTS

RESULT 1  
AAW80467  
ID AAW80467 standard; peptide; 16 AA.  
XX AC AAW80467;  
XX DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX DE Peptide derived from a conserved sequence of group O human HIV.  
XX KW Group O human immune deficiency virus; HIV; detection; infection.  
XX OS Synthetic.  
OS Human immunodeficiency virus 1.  
XX PN WO9845323-A1.  
XX PD 15-OCT-1998.  
XX PF 06-APR-1998; 98WO-FR000691.  
XX PR 09-APR-1997; 97FR-00004356.  
PR 24-FEB-1998; 98FR-00002212.  
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
PI Cheneboux DMB, Delagneau JH, Gadelle SJX, Rieunier PY;  
WPI; 1998-583190/49.  
DR New synthetic peptide(s) - useful for, e.g. detecting infection by human  
PT immune deficiency virus of group O.  
PT Claim 6; Page 43; 55pp; French.  
PS AAW80459-74 represent synthetic peptides (either linear or cyclised by  
CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
CC connected around short highly conserved sequences present in isolates of  
CC group O human immune deficiency virus (HIV). The peptides are useful as  
CC immunological reagents for detecting infection by group O human immune  
CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
XX (Updated on 27-AUG-2003 to correct OS field.)  
XX SQ Sequence 16 AA;  
SQ Query Match 100.0%; Score 91; DB 2; Length 16;

26	82	90.1	33	3 AAB12210	Aab12210 Partial s
27	82	90.1	33	3 AAB12221	Aab12221 Partial s
28	82	90.1	33	3 AAB12207	Aab12207 Partial s
29	82	90.1	33	3 AAB12213	Aab12213 Partial s
30	82	90.1	41	2 AAW07353	Aaw07353 Partial s
31	82	90.1	42	2 AAW07350	Aaw07350 Partial s
32	82	90.1	116	2 AAY05555	Aay05555 HIV-1 gro
33	82	90.1	117	2 AAY05548	Aay05548 HIV-1 gro
34	82	90.1	129	3 AAW69318	Aaw69318 Anti-HIV-
35	82	90.1	129	3 AAU77258	Aau77258 HIV-1 O-t
36	82	90.1	150	2 AAW69319	Aaw69319 Anti-HIV-
37	82	90.1	173	3 AAU77259	Aau77259 Protein A
38	82	90.1	200	2 AAY22908	Aay22908 SEQ ID NO
39	82	90.1	200	5 ABG68379	Abg68379 Envelope
40	82	90.1	200	6 ABU57786	Abu57786 Human imm
41	82	90.1	715	2 AAY05625	Aay05625 HIV-1 gro
42	81	89.0	23	2 AAY05615	Aay05615 HIV-1 gro
43	81	89.0	23	2 AAY05622	Aay05622 HIV-1 gro
44	81	89.0	23	2 AAY05619	Aay05619 HIV-1 gro
45	81	89.0	23	3 AAB12255	Aab12255 HIV-1 gp4



PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 XX WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX  
 XX Claim 6; Page 43; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 16 AA;  
 SQ  
 Query Match 96.7%; Score 88; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 3.4e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTS 16  
 DB 1 LLSSWGCKGRLVCYTS 16  
 II:|||||  
 RESULT 5  
 AAW80472  
 ID AAW80472 standard; peptide; 22 AA.  
 XX  
 AC AAW80472;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 PA  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX  
 XX Claim 6; Page 44; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 16 AA;  
 SQ  
 Query Match 96.7%; Score 88; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 3.4e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTS 16  
 DB 1 LLSSWGCKGRLVCYTS 16  
 II:|||||  
 RESULT 5  
 AAW80472  
 ID AAW80472 standard; peptide; 22 AA.  
 XX  
 AC AAW80472;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 PA  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX  
 XX Claim 6; Page 44; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 16 AA;  
 SQ  
 Query Match 96.7%; Score 88; DB 2; Length 22;  
 Best Local Similarity 93.8%; Pred. No. 4.6e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTS 16  
 DB 1 LLSSWGCKGRLVCYTS 16  
 II:|||||  
 RESULT 7  
 AAW80473  
 ID AAW80473 standard; peptide; 28 AA.  
 XX

CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 22 AA;  
 Query Match 96.7%; Score 88; DB 2; Length 22;  
 Best Local Similarity 93.8%; Pred. No. 4.6e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTS 16  
 DB 6 LLSSWGCKGRLVCYTS 21  
 II:|||||  
 RESULT 6  
 AAW80466  
 ID AAW80466 standard; peptide; 22 AA.  
 XX  
 AC AAW80466;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE Peptide derived from a conserved sequence of group O human HIV.  
 XX Group O human immune deficiency virus; HIV; detection; infection.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX  
 PN WO9845323-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 06-APR-1998; 98WO-FR000691.  
 XX  
 PR 09-APR-1997; 97FR-00004356.  
 PR 24-FEB-1998; 98FR-00002212.  
 XX  
 XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.  
 PA  
 XX Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;  
 PI WPI; 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.  
 XX  
 XX Claim 6; Page 43; 55pp; French.  
 PS  
 XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 22 AA;  
 SQ  
 Query Match 96.7%; Score 88; DB 2; Length 22;  
 Best Local Similarity 93.8%; Pred. No. 4.6e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRLVCYTS 16  
 DB 1 LLSSWGCKGRLVCYTS 16  
 II:|||||  
 RESULT 7  
 AAW80473  
 ID AAW80473 standard; peptide; 28 AA.  
 XX



XX Example 1; Fig 1; 52pp; English.

XX The present sequence is a partial gp41 protein of Human Immunodeficiency

CC Virus Type 1 (HIV-1) strain BCF13. HIV is the principle aetiological

CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV

CC envelope protein, and may be used as an antigen for the detection of

CC antibodies produced in response to HIV infection. BCF13 is a member of

CC HIV group O (outlier). The present sequence is the immunodominant region

CC of gp41. This sequence was used in a sequence homology alignment, which

CC in turn was used to derive a consensus sequence peptide; peptide 147

CC (AAB12254). (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 33 AA;

Query Match 96.7%; Score 88; DB 3; Length 33;

Best Local Similarity 93.8%; Pred. No. 6.8e-06;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16

DB 14 LLNSWGCKGRLVCYTS 29

RESULT 10

AAW07346

ID AAW07346 standard; peptide; 40 AA.

XX AC AAW07346;

XX DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF07 (MAN).

XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;

KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;

KW primer; hybridisation; amplification; PCR; polymerase chain reaction;

KW immunogen; antibody.

XX OS Human immunodeficiency virus 1.

XX PN WO9627013-A1.

XX PD 06-SEP-1996.

XX PF 26-FEB-1996; 96WO-FR000294.

XX PR 27-FEB-1995; 95PR-00002236.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX DR N-PSDB; AAT44922.

XX PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and

PT antibodies - useful for diagnosis, screening and typing, or as

PT immunogens.

XX PS Claim 12; Page 34; 71pp; French.

XX CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided

CC into 2 major groups based on the nucleotide sequences of the envelop gene

CC (env): group M containing sub-groups A-G, and group O containing the

CC strains ANT70 and MVP5180. The invention relates to the discovery of

CC several new strains of HIV-1 which can be placed in group O, based on the

CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39

CC and AAW07329-64). The novel strains have been deposited as retroviruses

CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547

CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from

CC the strain BCF13 and corresponds to a fragment of the gp41 protein

CC encoded by the env gene. The nucleic acids can be used to detect gp. O

CC HIV-1 strains by hybridisation or (as primers) by gene amplification,

CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547

CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from

CC the strain BCF07 (MAN) and corresponds to a fragment of the gp41 protein

CC encoded by the env gene. The nucleic acids can be used to detect gp. O

CC HIV-1 strains by hybridisation or (as primers) by gene amplification,

CC also for screening and typing of such strains. Peptides encoded by the

CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O

CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-

CC 2003 to standardise OS field)

XX SQ Sequence 40 AA;

Query Match 96.7%; Score 88; DB 2; Length 40;

Best Local Similarity 93.8%; Pred. No. 8.2e-06;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16

DB 19 LLNSWGCKGRLVCYTS 34

RESULT 11

AAW07352

ID AAW07352 standard; peptide; 40 AA.

XX AC AAW07352;

XX DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX DE Partial sequence of gp41 from HIV-1 gp. O strain BCF13.

XX KW Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;

KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;

KW primer; hybridisation; amplification; PCR; polymerase chain reaction;

KW immunogen; antibody.

XX OS Human immunodeficiency virus 1.

XX PN WO9627013-A1.

XX PD 06-SEP-1996.

XX PF 26-FEB-1996; 96WO-FR000294.

XX PR 27-FEB-1995; 95PR-00002236.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX PI Simon F, Saragosti S, Loussetajaka I, Ly T, Chaixbaudier M;

XX WPI; 1996-412779/41.

XX PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and

PT antibodies - useful for diagnosis, screening and typing, or as

PT immunogens.

XX PS Claim 12; Page 46; 71pp; French.

XX CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided

CC into 2 major groups based on the nucleotide sequences of the envelop gene

CC (env): group M containing sub-groups A-G, and group O containing the

CC strains ANT70 and MVP5180. The invention relates to the discovery of

CC several new strains of HIV-1 which can be placed in group O, based on the

CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39

CC and AAW07329-64). The novel strains have been deposited as retroviruses

CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547

CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from

CC the strain BCF13 and corresponds to a fragment of the gp41 protein

CC encoded by the env gene. The nucleic acids can be used to detect gp. O

CC HIV-1 strains by hybridisation or (as primers) by gene amplification,

CC also for screening and typing of such strains. Peptides encoded by the

CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O  
 CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)

XX Sequence 40 AA;

Query Match 96.7%; Score 88; DB 2; Length 40;  
 Best Local Similarity 93.8%; Pred. No. 8.2e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLSSWGCKGRVLCYTS 16  
 ||:|||||  
 Db 19 LLSSWGCKGRVLCYTS 34

RESULT 12

AAW80463  
 ID AAW80463 standard; peptide; 22 AA.

XX AC AAW80463;

XX DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JAN-1999 (first entry)

XX Peptide derived from a conserved sequence of group O human HIV.

XX Group O human immune deficiency virus; HIV; detection; infection.

OS Synthetic.

OS Human immunodeficiency virus 1.

XX WO9845323-A1.

XX PD 15-OCT-1998.

XX PF 06-APR-1998; 98WO-FR000691.

XX PR 09-APR-1997; 97FR-00004356.

PR 24-FEB-1998; 98FR-0002212.

XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.

XX Chenebaux DMB, Delagneau JH, Gadelle SUX, Rieunier FY;

XX WPI; 1998-583190/49.

XX New synthetic peptide(s) - useful for, e.g. detecting infection by human  
 PT immune deficiency virus of group O.

XX Claim 6; Page 42; 55pp; French.

XX AAW80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates of  
 CC group O human immune deficiency virus (HIV). The peptides are useful as  
 CC immunological reagents for detecting infection by group O human immune  
 CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 22 AA;

Query Match 95.6%; Score 87; DB 2; Length 22;  
 Best Local Similarity 93.8%; Pred. No. 6.5e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTS 16

||:|||||

Db 1 LLQSWGCKGRVLCYTS 16

RESULT 13

AAW12236

ID AAB12236 standard; peptide; 33 AA.

XX AC AAB12236;

XX DT 12-SEP-2003 (revised)

DT 10-NOV-2000 (first entry)

XX Partial sequence of HIV-1 strain ESS gp41 immunodominant region.

DE HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41; ESS.

XX Human immunodeficiency virus 1.

XX EP1013766-A2.

XX PD 28-JUN-2000.

XX PF 29-NOV-1999; 99EP-00309491.

XX PR 30-NOV-1998; 98US-0110292P.

PR 08-FEB-1999; 99US-0119138P.

PR 04-NOV-1999; 99US-00433428.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX De Leys R, Zheng J;

XX WPI; 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies.

PS Example 1; Fig 1; 52pp; English.

XX The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain ESS. HIV is the principle aetiological agent  
 CC for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope  
 CC protein, and may be used as an antigen for the detection of antibodies  
 CC produced in response to HIV infection. ESS is a member of HIV group O  
 CC (outlier). The present sequence is the immunodominant region of gp41.  
 CC This sequence was used in a sequence homology alignment, which in turn  
 CC was used to derive a consensus sequence peptide: peptide 147 (AAB12254).  
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 33 AA;

Query Match 94.5%; Score 86; DB 3; Length 33;  
 Best Local Similarity 87.5%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCYTS 16

||:|||||

Db 14 LLSSWGCKGRVLCYTS 29

RESULT 14

AAW07343

ID AAW07343 standard; peptide; 40 AA.

XX AC AAW07343;

XX DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-JUN-1997 (first entry)

XX Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (ESS).

XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;  
 KW C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;  
 KW primer; hybridisation; amplification; PCR; polymerase chain reaction;  
 KW immunogen; antibody.

```

XX OS Human immunodeficiency virus 1.
XX PN WO9627013-A1.
XX PD 06-SEP-1996.
XX PF 26-FEB-1996; 96WO-FR000294.
XX PR 27-FEB-1995; 95FR-00002236.
XX XX
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX PI Simon F, Saragosti S, Lousseertajaka I, Ly T, Chaixbaudier M;
XX DR WPI; 1996-412779/41.
XX DR N-PSDB; AAT44918.
XX XX
XX PT New strains of HIV-1 group O, related DNA fragments, peptide(s) and
XX PT antibodies - useful for diagnosis, screening and typing, or as
XX PT immunogens.
XX PS Claim 12; Page 33; 71pp; French.
XX XX
XX CC Human immunodeficiency virus type 1 (HIV-1) strains are currently divided
XX CC into 2 major groups based on the nucleotide sequences of the envelop gene
XX CC (env): group M containing sub-groups A-G, and group O containing the
XX CC strains ANT70 and MVP5180. The invention relates to the discovery of
XX CC several new strains of HIV-1 which can be placed in group O, based on the
XX CC partial sequences of the C2V3-env, gp41 and gag genes (see AAT44907-39
XX CC and AAW07329-64). The novel strains have been deposited as retroviruses
XX CC CNCM I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07 (MAN)), 1547
XX CC (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence presented here is from
XX CC the strain BCF02 (ESS) and corresponds to a fragment of the gp41 protein
XX CC encoded by the env gene. The nucleic acids can be used to detect gp. O
XX CC HIV-1 strains by hybridisation or (as primers) by gene amplification,
XX CC also for screening and typing of such strains. Peptides encoded by the
XX CC nucleic acids can be used as immunogens to raise Ab for detecting gp. O
XX CC HIV-1. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 16-OCT-
XX CC 2003 to standardise OS field)
XX SQ Sequence 40 AA;

Query Match 94.5%; Score 86; DB 2; Length 40;
Best Local Similarity 87.5%; Pred. No. 1.6e-05;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRIVCYTS 16
Db 19 LLNSWGCKGRIVCYTS 34

RESULT 15
AAW80469
ID AAW80469 standard; peptide; 32 AA.
XX AC
XX AC AAW80469;
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 28-JAN-1999 (first entry)
XX XX
XX DE Peptide derived from a conserved sequence of group O human HIV.
XX KW Group O human immune deficiency virus; HIV; detection; infection.
XX XX
XX OS Synthetic.
XX OS Human immunodeficiency virus 1.
XX PN WO9845323-A1.
XX PD 15-OCT-1998.

```

---

```

XX PF 06-APR-1998; 98WO-FR000691.
XX PR 09-APR-1997; 97FR-00004356.
XX PR 24-FEB-1998; 98FR-00002212.
XX XX
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX XX
XX PI Chenebaux DMB, Delagneau JH, Gadelle SJX, Rieunier FY;
XX DR WPI; 1998-583190/49.
XX XX
XX PT New synthetic peptide(s) - useful for, e.g. detecting infection by human
XX PT immune deficiency virus of group O.
XX PS Claim 6; Page 44; 55pp; French.
XX XX
XX CC AAW80459-74 represent synthetic peptides (either linear or cyclised by
XX CC Cys-Cys disulphide bonds). The peptides represent variable sequences
XX CC connected around short highly conserved sequences present in isolates of
XX CC group O human immune deficiency virus (HIV). The peptides are useful as
XX CC immunological reagents for detecting infection by group O human immune
XX CC deficiency virus (HIV). (Updated on 25-MAR-2003 to correct PI field.)
XX CC (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 32 AA;

Query Match 93.4%; Score 85; DB 2; Length 32;
Best Local Similarity 87.5%; Pred. No. 1.8e-05;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRIVCYTS 16
Db 11 LLNSWGCKGRIVCYTS 26

Search completed: May 7, 2004, 17:42:44
Job time : 31.4538 secs

```

THE UNIVERSITY OF CHICAGO



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:35:05 ; Search time 8.80672 Seconds  
(without alignments)  
93.794 Million cell updates/sec

Title: US-09-147-362A-9  
Perfect score: 91  
Sequence: 1 LLSSWGCKGRVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	96.7	33	3 US-09-433-428D-6	Sequence 6, Appli
2	88	96.7	33	3 US-09-433-428D-25	Sequence 25, Appl
3	88	96.7	40	3 US-08-894-699-39	Sequence 39, Appl
4	88	96.7	40	3 US-08-894-699-68	Sequence 68, Appl
5	88	96.7	40	3 US-09-444-410-39	Sequence 39, Appl
6	88	96.7	40	3 US-09-444-410-68	Sequence 68, Appl
7	86	94.5	33	3 US-09-433-428D-30	Sequence 30, Appl
8	86	94.5	40	3 US-08-894-699-36	Sequence 36, Appl
9	86	94.5	40	3 US-09-444-410-36	Sequence 36, Appl
10	85	93.4	33	3 US-09-433-428D-20	Sequence 20, Appl
11	85	93.4	40	3 US-08-894-699-41	Sequence 41, Appl
12	85	93.4	40	3 US-09-444-410-41	Sequence 41, Appl
13	82	90.1	23	3 US-09-433-428D-68	Sequence 68, Appl
14	82	90.1	23	4 US-09-462-917A-137	Sequence 137, App
15	82	90.1	24	4 US-09-462-917A-92	Sequence 92, Appl
16	82	90.1	24	4 US-09-462-917A-138	Sequence 138, App
17	82	90.1	32	4 US-09-625-972-30	Sequence 30, Appl
18	82	90.1	33	3 US-09-433-428D-1	Sequence 1, Appli
19	82	90.1	33	3 US-09-433-428D-4	Sequence 4, Appli
20	82	90.1	33	3 US-09-433-428D-7	Sequence 7, Appli
21	82	90.1	33	3 US-09-433-428D-15	Sequence 15, Appl
22	82	90.1	37	4 US-08-817-441-86	Sequence 86, Appl
23	82	90.1	41	3 US-08-894-699-69	Sequence 69, Appl
24	82	90.1	41	3 US-09-444-410-69	Sequence 69, Appl
25	82	90.1	42	3 US-08-894-699-66	Sequence 66, Appl
26	82	90.1	42	3 US-09-444-410-66	Sequence 66, Appl
27	82	90.1	116	4 US-09-462-917A-20	Sequence 20, Appl

28 90.1 117 4 US-09-462-917A-6 Sequence 6, Appli  
29 90.1 200 3 US-08-965-056-104 Sequence 104, App  
30 715 4 US-09-462-917A-134 Sequence 134, App  
31 862 4 US-09-206-551-15 Sequence 15, Appl  
32 81 89.0 23 3 US-09-433-428D-59 Sequence 59, Appl  
33 81 89.0 23 3 US-09-433-428D-60 Sequence 60, Appl  
34 81 89.0 23 3 US-09-433-428D-61 Sequence 61, Appl  
35 81 89.0 23 4 US-09-462-917A-95 Sequence 95, Appl  
36 81 89.0 23 4 US-09-462-917A-99 Sequence 99, Appl  
37 81 89.0 23 4 US-09-462-917A-102 Sequence 102, App  
38 81 89.0 28 3 US-09-433-428D-62 Sequence 62, Appl  
39 81 89.0 30 3 US-09-433-428D-63 Sequence 63, Appl  
40 81 89.0 30 3 US-09-433-428D-69 Sequence 69, Appl  
41 81 89.0 33 3 US-09-433-428D-5 Sequence 5, Appli  
42 81 89.0 33 3 US-09-433-428D-8 Sequence 8, Appli  
43 81 89.0 33 3 US-09-433-428D-14 Sequence 14, Appl  
44 81 89.0 33 3 US-09-433-428D-15 Sequence 15, Appl  
45 81 89.0 33 3 US-09-433-428D-29 Sequence 29, Appl

## ALIGNMENTS

RESULT 1  
US-09-433-428D-6  
; Sequence 6, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-6

Query Match 96.7%; Score 88; DB 3; Length 33;  
Best Local Similarity 93.8%; Pred. No. 6.7e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVCYTS 16  
Db 14 LLSSWGCKGRVCYTS 29

RESULT 2  
US-09-433-428D-25  
; Sequence 25, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-25

Query Match 96.7%; Score 88; DB 3; Length 33;  
Best Local Similarity 93.8%; Pred. No. 6.7e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 14 LLSWGCKGRLVCYTS 29

## RESULT 3

US-08-894-699-39  
; Sequence 39, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSET-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-39

Query Match 96.7%; Score 88; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 8.1e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 19 LLSWGCKGRLVCYTS 34

## RESULT 4

US-08-894-699-68  
; Sequence 68, Application US/08894699  
; Patent No. 6030769  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSET-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
; STREET: FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,699  
; FILING DATE: 01-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00294  
; FILING DATE: 26-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95/02236  
; FILING DATE: 27-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,614  
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-699-68

Query Match 96.7%; Score 88; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 8.1e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 19 LLSWGCKGRLVCYTS 34

## RESULT 5

US-09-444-410-39  
; Sequence 39, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:

; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSET-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.

STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
STREET: FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-39

Query Match 96.7%; Score 88; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 8.1e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVCYTS 16  
||:|||||||||  
Db 19 LLSWGCKGRVCYTS 34

RESULT 6  
US-09-444-410-68  
Sequence 68, Application US/09444410  
Patent No. 6270975  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-444-410-68

Query Match 96.7%; Score 88; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 8.1e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVCYTS 16  
||:|||||||||  
Db 19 LLSWGCKGRVCYTS 34

RESULT 7  
US-09-433-428D-30  
Sequence 30, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
APPLICANT: Zheng, Jian  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433,428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-30

Query Match 94.5%; Score 86; DB 3; Length 33;  
Best Local Similarity 87.5%; Pred. No. 1.3e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVCYTS 16  
||:|||||||||  
Db 14 LLSWGCKGRVCYTS 29

RESULT 8  
US-08-894-699-36  
Sequence 36, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
APPLICANT: LY, THOAI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE

```
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-699-36

Query Match 94.5%; Score 86; DB 3; Length 40;
Best Local Similarity 87.5%; Pred. No. 1.6e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRIVCYTS 16
Db 19 LLNSWGCKGRIVCYTS 34

RESULT 9
US-09-444-410-36
; Sequence 36, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; STREET: FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
```

```
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-444-410-36

Query Match 94.5%; Score 86; DB 3; Length 40;
Best Local Similarity 87.5%; Pred. No. 1.6e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRIVCYTS 16
Db 19 LLNSWGCKGRIVCYTS 34

RESULT 10
US-09-433-428D-20
; Sequence 20, Application US/09433428D
; Patent No. 6149910
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428D
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 33
; TYPE: PRI
; ORGANISM: Human immunodeficiency virus type 1
US-09-433-428D-20

Query Match 93.4%; Score 85; DB 3; Length 33;
Best Local Similarity 93.8%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSSWGCKGRIVCYTS 16
Db 14 LLSLWGCKGRIVCYTS 29

RESULT 11
US-08-894-699-41
; Sequence 41, Application US/08894699
; Patent No. 6030769
```

;; GENERAL INFORMATION:  
;; APPLICANT: SIMON, FRANCOIS  
;; APPLICANT: SARAGOSTI, SENTOB  
;; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
;; APPLICANT: LY, THOAI-DUONG  
;; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
;; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
;; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 81  
;; CORRESPONDENCE ADDRESS: 81  
;; ADDRESSEE: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; STREET: FLOOR  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/894,699  
;; FILING DATE: 01-DEC-1997  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR96/00294  
;; FILING DATE: 26-FEB-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614  
;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 41:  
;; FILING DATE: 27-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614  
;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 41:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-894-699-41

Query Match 93.4%; Score 85; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 2.3e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 19 LLSWGCKGRLVCYTS 34

RESULT 12  
US-09-444-410-41  
; Sequence 41, Application US/09444410  
; Patent No. 6270975  
; GENERAL INFORMATION:  
; APPLICANT: SIMON, FRANCOIS  
; APPLICANT: SARAGOSTI, SENTOB  
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM  
; APPLICANT: LY, THOAI-DUONG  
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS: 81  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

;; ADDRESSEE: P.C.  
;; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
;; STREET: FLOOR  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/444,410  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/894,699  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: FR 95/02236  
;; FILING DATE: 27-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,614  
;; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 41:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 40 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-09-444-410-41

Query Match 93.4%; Score 85; DB 3; Length 40;  
Best Local Similarity 93.8%; Pred. No. 2.3e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 19 LLSWGCKGRLVCYTS 34

RESULT 13  
US-09-433-428D-68  
; Sequence 68, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; APPLICANT: Zheng, Jian  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 68  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; US-09-433-428D-68

Query Match 90.1%; Score 82; DB 3; Length 23;  
Best Local Similarity 87.5%; Pred. No. 3.8e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
Db 4 LLSWGCKGRLVCYTS 19

Job time : 8.80672 secs

```

RESULT 14
US-09-462-917A-137
; Sequence 137, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NPUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 137
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
US-09-462-917A-137

```

```

Query Match          90.1%; Score 82; DB 4; Length 23;
Best Local Similarity 87.5%; Pred. No. 3.8e-06;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
Db 4 LLNLWGCKGRLVCYTS 19

```

```

RESULT 15
US-09-462-917A-92
; Sequence 92, Application US/09462917A
; Patent No. 6511801
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014 11362.0014.NPUS00
; CURRENT APPLICATION NUMBER: US/09/462,917A
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
US-09-462-917A-92

```

```

Query Match          90.1%; Score 82; DB 4; Length 24;
Best Local Similarity 87.5%; Pred. No. 4e-06;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
Db 4 LLNLWGCKGRLVCYTS 19

```

Search completed: May 7, 2004, 17:53:25

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:49:21 ; Search time 22.3866 Seconds  
(without alignments)  
198.381 Million cell updates/sec

Title: US-09-147-362A-9

Perfect score: 91  
Sequence: 1 LLSWGCKGRVLCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	90.1	23	14	US-10-320-786-137
2	82	90.1	24	14	US-10-320-786-92
3	82	90.1	24	14	US-10-320-786-138
4	82	90.1	32	14	US-10-364-360-30
5	82	90.1	37	14	US-10-026-741-86
6	82	90.1	116	14	US-10-320-786-20
7	82	90.1	117	14	US-10-320-786-6
8	82	90.1	200	9	US-09-854-816-104
9	82	90.1	715	14	US-10-320-786-134
10	82	90.1	862	14	US-10-369-294-15
11	81	89.0	23	14	US-10-320-786-95
12	81	89.0	23	14	US-10-320-786-99
13	81	89.0	23	14	US-10-320-786-102
14	81	89.0	110	14	US-10-320-786-38
15	81	89.0	113	14	US-10-320-786-28

16	81	89.0	113	14	US-10-320-786-40	Sequence 40, Appl
17	81	89.0	115	14	US-10-320-786-24	Sequence 24, Appl
18	81	89.0	215	8	US-08-911-824-58	Sequence 58, Appl
19	81	89.0	245	8	US-08-911-824-48	Sequence 48, Appl
20	81	89.0	281	8	US-08-911-824-120	Sequence 120, Appl
21	81	89.0	373	8	US-08-911-824-52	Sequence 52, Appl
22	81	89.0	460	8	US-08-911-824-60	Sequence 60, Appl
23	81	89.0	488	8	US-08-911-824-95	Sequence 95, Appl
24	81	89.0	490	8	US-08-911-824-50	Sequence 50, Appl
25	81	89.0	526	8	US-08-911-824-97	Sequence 97, Appl
26	81	89.0	618	8	US-08-911-824-54	Sequence 54, Appl
27	81	89.0	706	8	US-08-911-824-93	Sequence 93, Appl
28	81	89.0	736	8	US-08-911-824-91	Sequence 91, Appl
29	81	89.0	873	8	US-08-911-824-61	Sequence 61, Appl
30	80	87.9	23	14	US-10-320-786-91	Sequence 91, Appl
31	80	87.9	23	14	US-10-320-786-100	Sequence 100, Appl
32	80	87.9	23	14	US-10-320-786-101	Sequence 101, Appl
33	80	87.9	110	14	US-10-320-786-14	Sequence 14, Appl
34	80	87.9	110	14	US-10-320-786-16	Sequence 16, Appl
35	80	87.9	113	14	US-10-320-786-2	Sequence 2, Appl
36	80	87.9	113	14	US-10-320-786-10	Sequence 10, Appl
37	80	87.9	113	14	US-10-320-786-12	Sequence 12, Appl
38	80	87.9	113	14	US-10-320-786-18	Sequence 18, Appl
39	79	86.8	22	14	US-10-059-271-6	Sequence 6, Appl
40	79	86.8	22	14	US-10-059-271-8	Sequence 8, Appl
41	79	86.8	254	14	US-10-059-271-82	Sequence 82, Appl
42	79	86.8	256	14	US-10-059-271-97	Sequence 97, Appl
43	79	86.8	1231	14	US-10-059-271-94	Sequence 94, Appl
44	78	85.7	15	12	US-10-360-647A-8	Sequence 8, Appl
45	78	85.7	213	9	US-09-854-816-103	Sequence 103, Appl

#### ALIGNMENTS

RESULT 1  
US-10-320-786-137  
; Sequence 137, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362, 0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 137  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(23)  
; OTHER INFORMATION:  
US-10-320-786-137

Query Match 90.1%; Score 82; DB 14; Length 23;  
Best Local Similarity 87.5%; Pred. No. 3.8e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCYTS 16  
||:|||||  
Db 4 LLSWGCKGRVLCYTS 19

RESULT 2  
US-10-320-786-92  
; Sequence 92, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 92  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Human  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(24)  
; OTHER INFORMATION:  
US-10-320-786-92

Query Match 90.1%; Score 82; DB 14; Length 24;  
Best Local Similarity 87.5%; Pred. No. 4e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
||: |||||

Db 4 LLSWGCKGRLVCYTS 19

RESULT 3  
US-10-320-786-138  
; Sequence 138, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 138  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Human  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(24)  
; OTHER INFORMATION:  
US-10-320-786-138

Query Match 90.1%; Score 82; DB 14; Length 24;  
Best Local Similarity 87.5%; Pred. No. 4e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
||: |||||

Db 4 LLSWGCKGRLVCYTS 19

## RESULT 4

US-10-364-360-30  
; Sequence 30, Application US/10364360  
; Publication No. US20030180324A1  
; GENERAL INFORMATION:  
; APPLICANT: GUERTLER, Lutz Gerhard  
; APPLICANT: HAUSER, Hans Peter  
; APPLICANT: DONGMO DELOKO, Yvette Beatrice  
; APPLICANT: ZEKENG, Leopold  
; APPLICANT: KAPTUE, Lazare  
; TITLE OF INVENTION: LENTIVIRUS FROM THE GROUP OF IMMUNODEFICIENCY VIRUSES OF DRILL MONKEY  
; FILE REFERENCE: 067595/0106  
; CURRENT APPLICATION NUMBER: US/10/364,360  
; CURRENT FILING DATE: 2003-02-12  
; PRIOR APPLICATION NUMBER: US/09/625,972  
; PRIOR FILING DATE: 2000-07-29  
; PRIOR APPLICATION NUMBER: DE 199 36 003.0  
; PRIOR FILING DATE: 1999-08-03  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 30  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: HIV1-O, ANT70C  
US-10-364-360-30

Query Match 90.1%; Score 82; DB 14; Length 32;  
Best Local Similarity 87.5%; Pred. No. 5.1e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16  
||: |||||

Db 14 LLSWGCKGRLVCYTS 29

## RESULT 5

US-10-026-741-86  
; Sequence 86, Application US/10026741  
; Publication No. US20030049604A1  
; GENERAL INFORMATION:  
; APPLICANT: CHARNEAU, PIERRE  
; CLAVEL, FRANCOISE  
; BORMAN, ANDREW  
; QUILLENT, CAROLINE  
; GUETARD, DENISE  
; MONTAGNIER, LUC  
; DONJON DE SAINT-MARTIN, JACQUELINE  
; COHEN, JAKUES  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR SUBTYPE) ANTIGENS  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/026,741  
; FILING DATE: 27-Dec-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/817,441



;  
; FILING DATE: 31-AUG-1998  
; APPLICATION NUMBER: PCT/FR 95/01391  
; FILING DATE: 20-OCT-1995  
; APPLICATION NUMBER: FR 9412554  
; FILING DATE: 20-OCT-1994  
; APPLICATION NUMBER: FR 9502526  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03260.6005-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
US-10-026-741-86

Query Match 90.1%; Score 82; DB 14; Length 37;  
Best Local Similarity 87.5%; Pred. No. 5.9e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
||| ||||| |||||  
Db 14 LLSLWGCKGRLVCYTS 29

RESULT 6  
US-10-320-786-20  
; Sequence 20, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014-1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-20

Query Match 90.1%; Score 82; DB 14; Length 116;  
Best Local Similarity 87.5%; Pred. No. 0.00016;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
||| ||||| |||||  
Db 39 LLMLWGCKGRLVCYTS 54

RESULT 7  
US-10-320-786-6  
; Sequence 6, Application US/10320786  
; Publication No. US20030180759A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaporte, Eric

;  
; APPLICANT: Peeters, Martine  
; APPLICANT: Saman, Eric  
; APPLICANT: Vanden Haesevelde, Marlen  
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof  
; FILE REFERENCE: INNS:014-1 11362.0014.DVUS01  
; CURRENT APPLICATION NUMBER: US/10/320,786  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: US 09/462,917  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/EP98/04522  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 152  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Human  
US-10-320-786-6

Query Match 90.1%; Score 82; DB 14; Length 117;  
Best Local Similarity 87.5%; Pred. No. 0.00016;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
||| ||||| |||||  
Db 39 LLMLWGCKGRLVCYTS 54

RESULT 8  
US-09-854-816-104  
; Sequence 104, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell  
; J. Christopher Phelan  
; Melissa A. Starovaenik  
; James A. Wells  
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/854,816  
; FILING DATE: 15-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/965,056  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, Ph.D., Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: P1005R2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-8674  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 104:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 200 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear

```
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-854-816-104

Query Match          90.1%; Score 82; DB 9; Length 200;
Best Local Similarity 87.5%; Pred. No. 0.00027;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16
   ||| |||||:|||||
Db 126 LLSWGCKGRLVCYTS 141

RESULT 9
US-10-320-786-134
; Sequence 134, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-134

Query Match          90.1%; Score 82; DB 14; Length 715;
Best Local Similarity 87.5%; Pred. No. 0.00083;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16
   ||| |||||:|||||
Db 628 LLSWGCKGRLVCYTS 643

RESULT 10
US-10-369-294-15
; Sequence 15, Application US/10369294
; Publication No. US20030162170A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice H.
; APPLICANT: Gao, Feng
; APPLICANT: Marx, Preston A.
; APPLICANT: Shaw, George M.
; APPLICANT: Smith, Stephen M.
; APPLICANT: Georges-Courbot, Marie Claude
; APPLICANT: Lu, Chang Yong
; TITLE OF INVENTION: Complete Genome Sequences of a Simian
; TITLE OF INVENTION: Immunodeficiency Virus from a Red-Capped
; TITLE OF INVENTION: Mangabey
; FILE REFERENCE: D6286D
; CURRENT APPLICATION NUMBER: US/10/369,294
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 09/206,551
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 58
; SEQ ID NO 15
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of homologous region of

;
; OTHER INFORMATION: O_ANT70 lentiviral env protein
US-10-369-294-15

Query Match          90.1%; Score 82; DB 14; Length 862;
Best Local Similarity 87.5%; Pred. No. 0.00098;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16
   ||| |||||:|||||
Db 593 LLSWGCKGRLVCYTS 608

RESULT 11
US-10-320-786-95
; Sequence 95, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-95

Query Match          89.0%; Score 81; DB 14; Length 23;
Best Local Similarity 81.2%; Pred. No. 5.4e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLVCYTS 16
   ||| |||||:|||||
Db 4 LLSWGCKGRLVCYTS 19

RESULT 12
US-10-320-786-99
; Sequence 99, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
US-10-320-786-99
```

```
;
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-99

Query Match      89.0%; Score 81; DB 14; Length 23;
Best Local Similarity 81.2%; Pred. No. 5.4e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
   ||| |||||:||||
Db 4 LLNLWGCKGRLICYTS 19

RESULT 13
US-10-320-786-102
; Sequence 102, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-102

Query Match      89.0%; Score 81; DB 14; Length 23;
Best Local Similarity 81.2%; Pred. No. 5.4e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
   ||| |||||:||||
Db 4 LLNLWGCKGRLICYTS 19

RESULT 14
US-10-320-786-38
; Sequence 38, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38

;
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-38

Query Match      89.0%; Score 81; DB 14; Length 110;
Best Local Similarity 81.2%; Pred. No. 0.00022;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
   ||| |||||:||||
Db 39 LLNLWGCKGRLICYTS 54

RESULT 15
US-10-320-786-28
; Sequence 28, Application US/10320786
; Publication No. US20030180759A1
; GENERAL INFORMATION:
; APPLICANT: Delaporte, Eric
; APPLICANT: Peeters, Martine
; APPLICANT: Saman, Eric
; APPLICANT: Vanden Haesevelde, Marlen
; TITLE OF INVENTION: HIV-1 Group O Antigens and Uses Thereof
; FILE REFERENCE: INNS:014--1 11362.0014.DVUS01
; CURRENT APPLICATION NUMBER: US/10/320,786
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 09/462,917
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/EP98/04522
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-10-320-786-28

Query Match      89.0%; Score 81; DB 14; Length 113;
Best Local Similarity 81.2%; Pred. No. 0.00022;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
   ||| |||||:||||
Db 39 LLNLWGCKGRLICYTS 54

Search completed: May 7, 2004, 18:29:21
Job time : 22.3866 secs
```

Western Nevada Electric Co.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:34:00 ; Search time 6.72269 Seconds  
(without alignments)  
228.936 Million cell updates/sec

Title: US-09-147-362A-9  
Perfect score: 91  
Sequence: 1 LLSSWCKGKRLVCYTS 16  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	90.1	863	2 A53034	gag polyprotein -
2	75	82.4	877	2 S49197	envelope protein p
3	70	76.9	104	2 S52930	GP41 ENV protein -
4	63	69.2	357	2 S21990	envelope protein g
5	63	69.2	358	2 S22002	envelope protein g
6	63	69.2	854	1 VCLJST	env polyprotein pr
7	62	68.1	358	2 S22000	envelope protein g
8	62	68.1	358	2 S70417	envelope protein g
9	60	65.9	357	2 S22006	envelope protein g
10	60	65.9	357	2 S21994	envelope protein g
11	60	65.9	357	2 S22004	envelope protein g
12	60	65.9	357	2 S21996	envelope protein g
13	60	65.9	357	2 S21992	envelope protein g
14	60	65.9	358	2 S21998	envelope protein g
15	60	65.9	443	2 C41621	env polyprotein p
16	60	65.9	445	2 A41621	env polyprotein M
17	60	65.9	454	2 B41621	env polyprotein D
18	60	65.9	843	1 H40401	env polyprotein pr
19	60	65.9	847	2 T09448	envelope glycoprot
20	60	65.9	847	2 S13289	env protein - huma
21	60	65.9	852	1 VCLJBR	env polyprotein -
22	60	65.9	852	2 T12016	envelope glycoprot
23	60	65.9	853	2 S54384	envelope polyprote
24	60	65.9	854	2 S13288	env protein - huma
25	60	65.9	855	1 VCLJAZ2	env polyprotein pr
26	60	65.9	855	1 VCLJZR	env polyprotein pr
27	60	65.9	856	1 VCLJH3	env polyprotein pr
28	60	65.9	856	1 VCLJVL	env polyprotein pr
29	60	65.9	856	1 VCLJ3W	env polyprotein pr

30	60	65.9	859	1 VCLJMN	env polyprotein pr
31	60	65.9	861	1 VCLJLV	env polyprotein pr
32	60	65.9	861	1 VCLJSC	env polyprotein pr
33	60	65.9	868	1 VCLJH4	env polyprotein -
34	59	64.8	855	2 A45713	Env transmembrane
35	58	63.7	859	1 VCLJST	env polyprotein pr
36	57	62.6	786	2 S28084	env polyprotein -
37	57	62.6	856	1 A44963	env polyprotein pr
38	56	61.5	846	1 VCLJND	env polyprotein pr
39	55	60.4	151	2 S30458	env protein - huma
40	55	60.4	151	2 S30459	env protein - huma
41	55	60.4	151	2 S30448	env protein - huma
42	55	60.4	151	2 S30453	env protein - huma
43	55	60.4	151	2 S30452	env protein - huma
44	55	60.4	151	2 S30450	env protein - huma
45	55	60.4	151	2 S30451	env protein - huma

ALIGNMENTS

RESULT 1

A53034

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Lays, R.J.; Vanderborght, B.; van der Groen, J. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African h

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-863 <VAN>

A:Cross-references: GB:I02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 90.1%; Score 82; DB 2; Length 863;  
Best Local Similarity 87.5%; Pred. No. 8.6e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSSWCKGKRLVCYTS 16

||| |||||:|||||

Db 594 LLSSWCKGKRLVCYTS 609

RESULT 2

S49197

envelope protein precursor - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 24-Jul-1997 #text\_change 26-Aug-1999

C:Accession: S49197

R:Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Chamaret, S.; Cohen, J.; Remy, submitted to the EMBL Data Library, July 1994

A:Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate: def

A:Reference number: S49197

A:Accession: S49197

A:Molecule type: DNA

A:Residues: 1-877 <CHA>

A:Cross-references: EMBL:X80020; NID:gs10516; PIDN:CAAS6323.1; PID:gs10517

A:Experimental source: isolate VAU

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-535/Product: coat protein gp120 #status predicted <CP1>

F:536-877/Product: coat protein gp41 #status predicted <CP2>

F:698-716/Domain: transmembrane #status predicted <TMN>

F:59,88,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404

Query Match 82.4%; Score 75; DB 2; Length 877;  
Best Local Similarity 75.0%; Pred. No. 0.001;

```
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRRLVCYTS 16
   ||: |||| ||: ||||
Db 607 LLNLWGCKRRLICYTS 622

RESULT 3
S22930
Gp41 ENV protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Jun-1995 #sequence_revision 21-Jul-1995 #text_change 26-Aug-1999
C:Accession: S52930
R:Cohen, J.H.M.; Guetard, D.; Philibert, F.; Chamaret, S.; Tabary, T.; Montagnier, L.; de
submitted to the EMBL Data Library, January 1995
A:Description: A novel HIV-1 O strain illustrates the diversity of the O group.
A:Reference number: S52929
A:Accession: S52930
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-104 <COH>
A:Cross-references: EMBL:X84328; NID:G95526; PIDN:CAA59066.1; PID:G695527
C:Superfamily: type E retrovirus env polyprotein

Query Match 76.9%; Score 70; DB 2; Length 104;
Best Local Similarity 62.5%; Pred. No. 0.0011; 2; Indels 0; Gaps 0;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRRLVCYTS 16
   ||: |||| ||: ||||
Db 45 LLNLWGCKRRLICYTS 60

RESULT 4
S21990
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 20
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21990; S70423
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S21990
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <STEL>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70423
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332, 'X', 334-357 <STE2>
A:Cross-references: EMBL:X61357; NID:g60175; PIDN:CAA43626.1; PID:g60176
C:Superfamily: type E retrovirus env polyprotein

Query Match 69.2%; Score 63; DB 2; Length 357;
Best Local Similarity 62.5%; Pred. No. 0.034;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRRLVCYTS 16
   || ||| ||||: |||
Db 93 LLGIWGCSGRRLICTTA 108

RESULT 5
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 3L
```

```
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21990
A:Accession: S22002
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STEL>
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebri
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, 'X', 335-358 <STE2>
A:Cross-references: EMBL:X61352; NID:g60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 69.2%; Score 63; DB 2; Length 358;
Best Local Similarity 62.5%; Pred. No. 0.034;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRRLVCYTS 16
   || ||| ||||: |||
Db 94 LLGIWGCSGRRLICTTA 109

RESULT 6
VCLJSI
env polyprotein precursor - simian immunodeficiency virus SIVcpz
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: simian immunodeficiency virus SIVcpz
A:Note: host Pan troglodytes (Chimpanzee)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09990
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09990
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-854 <HUE>
A:Cross-references: EMBL:X52154; NID:G58866; PIDN:CAA36407.1; PID:g59874
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-500/Product: coat protein gp120 #status predicted <CP1>
F:501-854/Product: coat protein gp41 #status predicted <CP2>
F:501-517/Domain: transmembrane #status predicted <TM1>
F:675-693/Domain: transmembrane #status predicted <TM2>
F:805-821/Domain: transmembrane #status predicted <TM3>
F:134,140,143,154,158,186,195,239,267,274,299,331,336,351,356,384,392,426,432,446,45

Query Match 69.2%; Score 63; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 0.068;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRRLVCYTS 16
   || ||| ||||: |||
Db 582 ILGLWGCSGRKAVCYTT 597

RESULT 7
S22000
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
```

```
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 12-Apr-1995
C;Accession: S22000
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined
A;Reference number: S21990
A;Accession: S22000
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <STE>
A;Cross-references: EMBL:X61351
C;Superfamily: type E retrovirus env polyprotein

Query Match      68.1%; Score 62; DB 2; Length 358;
Best Local Similarity 66.7%; Pred. No. 0.048;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYT 15
Db 94 LLGIWCGSGRLICTT 108

RESULT 8
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 3B) (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: patient 3B
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 26-Aug-1999
C;Accession: S70417
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70417
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <STE>
A;Cross-references: EMBL:X61351; NID:g60184; PIDN:CAA43614.1; PID:g60185
C;Superfamily: type E retrovirus env polyprotein

Query Match      68.1%; Score 62; DB 2; Length 358;
Best Local Similarity 66.7%; Pred. No. 0.048;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYT 15
Db 94 LLGIWCGSGRLICTT 108

RESULT 9
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S70420; S22006
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <STE>
A;Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191
A;Experimental source: patient L
A;Note: submitted to the EMBL Data Library, July 1991
C;Superfamily: type E retrovirus env polyprotein

Query Match      65.9%; Score 60; DB 2; Length 357;
Best Local Similarity 56.2%; Pred. No. 0.097;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
Db 93 LLGIWCGSGRLICTTT 108

RESULT 10
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 27B
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S21994; S70421
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined
A;Reference number: S21990
A;Accession: S21994
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43622.1; PID:g60180
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140,'X',142-312,'X',314-357 <STE2>
A;Cross-references: EMBL:X61355; NID:g60179
C;Superfamily: type E retrovirus env polyprotein

Query Match      65.9%; Score 60; DB 2; Length 357;
Best Local Similarity 56.2%; Pred. No. 0.097;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
Db 93 LLGIWCGSGRLICTTT 108

RESULT 11
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 4B
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C;Accession: S22004; S70419
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined
A;Reference number: S21990
A;Accession: S22004
A;Molecule type: DNA
A;Residues: 1-357 <STE1>
A;Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid as determined
A;Reference number: S70417; MUID:92144209; PMID:1736940
A;Accession: S70419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-292,'X',294-357 <STE2>
A;Cross-references: EMBL:X61353; NID:g60188
C;Superfamily: type E retrovirus env polyprotein

Query Match      65.9%; Score 60; DB 2; Length 357;
Best Local Similarity 56.2%; Pred. No. 0.097;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16
Db 93 LLGIWCGSGRLICTTT 108
```

RESULT 12

S21996  
 envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
 C:Accession: S70422; S21996  
 R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
 AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
 A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
 A:Reference number: S70417; MUID:92144209; PMID:1736940  
 A:Accession: S70422  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-357 <ST2>  
 A:Cross-references: EMBL:X61356; NID:G60181; PIDN:CAA43624.1; PID:g1067129  
 A:Experimental source: patient 27L  
 A:Note: submitted to the EMBL Data Library, July 1991  
 C:Superfamily: type E retrovirus env polyprotein

```

Query Match      65.9%; Score 60; DB 2; Length 357;
Best Local Similarity 56.2%; Pred. No. 0.097;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRIVCVTS 16
    |||||.:|:
Db 93 LLGIWGCKGICTTA 108

```

RESULT 13

S21992

envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 22)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S70424; S21992

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70424

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-357 <S72>

A:Cross-references: EMBL:X61358; NID:g60177; PIDN:CAA43628.1; PID:g60178

A:Experimental source: patient 22

A:Note: Submitted to the EMBL Data Library, July 1991

C:Superfamily: type E retrovirus env polyprotein

```
Query Match      65.9%; Score 60; DB 2; Length 357;
Best Local Similarity 56.2%; Pred.No. 0.097;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

RESULT 14

S21998  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Variety: isolate 28  
C;date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C;Accession: S21998; S70425  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A;description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined by polymerase chain reaction  
A;Reference number: S21990  
A;Accession: S21998  
A;Molecule type: DNA  
A;Residues: 1-358 <STE1>  
A;Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183  
R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A>Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70425  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-222,'X',224-358 <STE2>  
A:Cross-references: ENBL:X61359; NID:g60182; PID:g60183  
C:Superfamily: type E retrovirus env polypeptide

```

Query Match      65.9%; Score 60; DB 2; Length 358;
Best Local Similarity 56.4%; Pred. No. 0.098;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLSWGCKGRIVCYTS 16
    |||||.:|:|
Db 94 LLGWGCGRLICTTT 109

```

```

RESULT 15
C41621
env polypeptide P - human immunodeficiency virus type 1 (fragment)
N:Alternate names: coat polypeptide
N:Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 11-Feb-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C:Accession: C41621
R:Burger, H.; Weiser, B.; Flaherty, K.; Gulla, J.; Nguyen, P.N.; Gibbs, R.A.
Proc.Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991
A:Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A:Reference number: A41621; MUID:92107924; PMID:1763038
A:Accession: C41621
A:Molecule type: DNA
A:Residues: 1-443 <BUR>
A:Cross-references: GB:M77230; NID:g328631; PIDN:AAB03792.1; PID:g555015
A:Note: This virus was isolated from the mother's sexual partner
C:Genetics:
A:Gene: env

```

```

Query Match      65.9%; Score 60; DB 2; Length 443;
Best Local Similarity 56.2%; Pred. NO. 0.12;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      1  LLSSWGCKGRIVCVYTS 16
      ||| ||| ||| ||| |||
Db      332  LKLGWCSGKLICTTA 347

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:32:55 ; Search time 4.03361 Seconds  
(without alignments)  
206.545 Million cell updates/sec

Title: US-09-147-362A-9  
Perfect score: 91  
Sequence: 1 L1SSWGCKGRLVCYTS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141691 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	69.2	854	1 ENV_SIVCZ	P17281 chimpanzee
2	60	65.9	843	1 ENV_HV1Y2	P35961 human immun
3	60	65.9	847	1 ENV_HV1S1	P19550 human immun
4	60	65.9	847	1 ENV_HV1W2	P05880 human immun
5	60	65.9	851	1 ENV_HV1B8	P04582 human immun
6	60	65.9	852	1 ENV_HV1BN	P12488 human immun
7	60	65.9	852	1 ENV_HV1S3	P19549 human immun
8	60	65.9	853	1 ENV_HV1MF	P19551 human immun
9	60	65.9	853	1 ENV_HV1Z2	P12487 human immun
10	60	65.9	855	1 ENV_HV1A2	P03378 human immun
11	60	65.9	855	1 ENV_HV1OY	P20888 human immun
12	60	65.9	855	1 ENV_HV1Z6	P04580 human immun
13	60	65.9	856	1 ENV_HV1B1	P03375 human immun
14	60	65.9	856	1 ENV_HV1H2	P04578 human immun
15	60	65.9	856	1 ENV_HV1LW	P070626 human immun
16	60	65.9	856	1 ENV_HV1MN	P05877 human immun
17	60	65.9	856	1 ENV_HV1PV	P03376 human immun
18	60	65.9	856	1 ENV_HV1SC	P05878 human immun
19	60	65.9	856	1 ENV_HV1W1	P31872 human immun
20	60	65.9	857	1 ENV_HV2KR	P04126 human immun
21	60	65.9	861	1 ENV_HV1BR	P03377 human immun
22	60	65.9	865	1 ENV_HV1RH	P04579 human immun
23	60	65.9	867	1 ENV_HV1J3	P12489 human immun
24	60	65.9	868	1 ENV_HV1C4	P05879 human immun
25	58	63.7	848	1 ENV_HV1JR	P20871 human immun
26	58	63.7	856	1 ENV_HV1H3	P04624 human immun
27	58	63.7	856	1 ENV_HV2N2	P05883 human immun
28	57	62.6	821	1 ENV_SIVGB	P22380 simian immu
29	57	62.6	856	1 ENV_HV1ZH	P05881 human immun
30	56	61.5	846	1 ENV_HV1ND	P18799 human immun
31	55	60.4	712	1 ENV_HV2S2	P32536 human immun
32	55	60.4	846	1 ENV_HV2SB	P12449 human immun
33	55	60.4	851	1 ENV_HV2D1	P17755 human immun

#### ALIGNMENTS

RESULT 1

ID	ENV_SIVCZ	STANDARD;	PRT;	854 AA.
DT	P17281;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
DE	ENV.			
GN	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retrovirus; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyerhans A., Roelants G., Main-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."			
RC	Nature 345:356-359(1990).			
CC	-I- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.			
CC	This SWISS-PROT entry is copyright: It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; X52154; CA36407.1; -			
DR	PIR; S09990; VCLJ81.			
DR	HIV; X52154; ENVSCPZ.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .) (POTENTIAL).

34	55	60.4	851	1	ENV_HV2G1	P18040 human immun
35	55	60.4	858	1	ENV_HV2RO	P04577 human immun
36	55	60.4	859	1	ENV_HV2CA	P24105 human immun
37	55	60.4	859	1	ENV_HV2D2	P15831 human immun
38	55	60.4	859	1	ENV_HV2ST	P20872 human immun
39	55	60.4	860	1	ENV_HV2BE	P18094 human immun
40	55	60.4	885	1	ENV_SIVS4	P12492 simian immu
41	55	60.4	889	1	ENV_SIVSP	P19503 simian immu
42	54	59.3	861	1	ENV_HV1KB	P31819 human immun
43	53	58.2	853	1	ENV_HV1EL	P04581 human immun
44	53	58.2	854	1	ENV_SIVAI	Q02837 simian immu
45	53	58.2	859	1	ENV_HV1MA	P04583 human immun

```
FT CARBOHYD 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 69.2%; Score 63; DB 1; Length 854;
Best Local Similarity 56.2%; Pred. No. 0.013;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLYCYTS 16
DB 582 ILGLMGCSGKAVCYTT 597

RESULT 2
ENV_HV1Y2
ID ENV_HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retrov. Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
properties of human immunodeficiency virus type 1 in vivo: evidence
for limited detectiveness and complementation."
RL J. Virol. 66:6587-6600(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
CC FIR; H44001; H44001.
CC PDB; 1G9N; 27-DEC-00.
CC InterPro; IPR000328; Env GP41.
CC DR InterPro; IPR000777; GP120.
CC DR Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
KW SIGNAL 1 29
FT CHAIN 30 499 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
```

```
FT DISULFID 214 BY SIMILARITY.
FT DISULFID 224 BY SIMILARITY.
FT DISULFID 292 BY SIMILARITY.
FT DISULFID 373 BY SIMILARITY.
FT DISULFID 380 BY SIMILARITY.
FT CARBOHYD 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 193 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 381 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 389 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 803 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 65.9%; Score 60; DB 1; Length 843;
Best Local Similarity 56.2%; Pred. No. 0.039;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSWGCKGRLYCYTS 16
DB 579 LLGIWGCSGKLICTTT 594

RESULT 3
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retrov. Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation."
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC CC
```



```

Db      583 LLGIWCSGKLICTTT 598
RESULT 5
ID ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (RH8 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02011; AAA44661.1; -.
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QO3; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENV$B8.
DR GlycoSuiteDB; P04582; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 506
FT CHAIN 507 851
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 236 331
FT DISULFID 378 440
FT DISULFID 385 413
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 401 401
FT CARBOHYD 443 443
FT CARBOHYD 458 458
FT CARBOHYD 606 606
FT CARBOHYD 611 611
FT CARBOHYD 620 620
FT CARBOHYD 632 632
FT CARBOHYD 669 669
FT CARBOHYD 745 745
FT CARBOHYD 811 811
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;
Query Match 65.9%; Score 60; DB 1; Length 851;
Best Local Similarity 56.2%; Pred. No. 0.04;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QV 1 LLSSWGCKGRLVCYTS 16
||| ||||| :||| :|
Db 587 LLGIWCSGKLICTTA 602
RESULT 6
ENV_HV1B8
ID ENV_HV1B8 STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -!- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21098; AAA44221.1; -.
DR PIR; A31667; VCLJBR.
DR PDB; 1IM7; 23-OCT-02.
DR HIV; M21098; ENV$BRVA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 507
FT CHAIN 508 852
FT CHAIN 853 852
FT EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEMBRANE GLYCOPROTEIN.

```





```

FT CARBOHYD 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

Query Match 65.9%; Score 60; DB 1; Length 853;
Best Local Similarity 56.2%; Pred. No. 0.04;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCVTS 16
Db 589 LLGIWCSGKLICTTT 604

RESULT 10
ENV_HV1A2
ID ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
(RARV-2).";
RL Science 227:484-492(1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; K02007; AAB59882.1; -.
DR FIC; A03976; VCLJJA2.
DR HIV; K02007; ENV5SF2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
Signal.
FT CHAIN 1 29 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 30 509 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 510 855 BY SIMILARITY.
FT DISULFID 53 73
FT DISULFID 118 208

```

```

FT DISULFID 125 199 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 221 250 BY SIMILARITY.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 299 333 BY SIMILARITY.
FT DISULFID 380 442 BY SIMILARITY.
FT DISULFID 387 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

Query Match 65.9%; Score 60; DB 1; Length 855;
Best Local Similarity 56.2%; Pred. No. 0.04;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRVLCVTS 16
Db 591 LLGIWCSGKLICTTA 606

RESULT 11
ENV_HV10Y
ID ENV_HV10Y STANDARD; PRT; 855 AA.
AC P20888;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
HEALTHY GABONESE INDIVIDUAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

DR	ENBL; M26727; AA83397.1; -.	
DR	HIV; M26727; ENVSOYI.	
DR	InterPro; IPR000328; Envv GP41.	
DR	InterPro; IPR000777; GP120.	
DR	Pfam; PF00516; GP120; 1.	
DR	Pfam; PF00517; GP41; 1.	
DR	ADAMS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.	
KW	FT SIGNAL.	1 29

RESULT 12	
ENV_HV126	
ID ENV_HV126	STANDARD;
AC P04580:	PRT; 855 AA.
DT 13-AUG-1987	(Rel. 05, Created)
DT 13-AUG-1987	(Rel. 05, Last sequence update)
DT 15-JUL-1999	(Rel. 38, Last annotation update)
DE Envelope polypeptide Gp160 precursor	(Contains: Exterior membrane





-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; K03455; AAB50262.1; -;  
EMBL; AF038399; AAB99976.1; -;  
PDB; 1DF4; 26-JAN-00.  
PDB; 1DF5; 26-JAN-00.  
PDB; 1DLB; 02-APR-00.  
PDB; 1G9M; 27-DEC-00.  
PDB; 1GCL; 19-AUG-98.  
PDB; 1GZL; 10-OCT-02.  
PDB; 1K33; 10-OCT-01.  
PDB; 1K34; 10-OCT-01.  
HIV; K03455; ENVSHXB2.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
3D-structure.

FT	SIGNAL	1	30	
FT	CHAIN	31	511	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	512	856	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	BY SIMILARITY.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	236	331	BY SIMILARITY.
FT	DISULFID	378	445	BY SIMILARITY.
FT	DISULFID	385	418	BY SIMILARITY.
FT	CARBOHYD	88		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	136		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	141		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	156		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	160		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	186		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	197		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	230		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	234		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	241		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	262		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	276		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	289		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	295		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	301		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	332		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	339		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	356		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	356		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	386		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	392		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	397		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	406		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	448		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	463		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	611		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	616		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	624		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	637		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	674		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	750		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	816		N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	856 AA;	97212 MW;	6FAB16AF85107FE0 CRC64;

Query Match

65.9%; Score 60; DB 1; Length 856;

Best Local Similarity 56.2%; Pred. No. 0.04;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 LLSSWGCKGRIVCYTS 16  
Db 592 LLGIWCGSGKLICTTA 607  
RESULT 15  
ENV\_HVILW  
ID ENV\_HVILW STANDARD; PRT; 856 AA.  
AC Q70626;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=82834;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95127297; PubMed=7826699;  
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M., Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HTLV type IIIB)."  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U12055; AAA76690.1; -;  
PDB; 1IF3; 02-MAY-01.  
DR GlycoSuiteDB; Q70626; -;  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
KW 3D-structure.  
FT SIGNAL 1 30  
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 205 BY SIMILARITY.  
FT DISULFID 126 196 BY SIMILARITY.  
FT DISULFID 131 157 BY SIMILARITY.  
FT DISULFID 218 247 BY SIMILARITY.  
FT DISULFID 228 239 BY SIMILARITY.  
FT DISULFID 296 331 BY SIMILARITY.  
FT DISULFID 378 445 BY SIMILARITY.  
FT DISULFID 385 418 BY SIMILARITY.  
FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 406 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 448 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 637 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 674 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 750 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	448	448	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	674	674	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	750	750	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	816	816	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	856 AA;	96938 MW;	OC241332CF7E6687	CRC64;

Query Match 65.9%; Score 60; DB 1; Length 856;  
Best Local Similarity 56.2%; Pred. No. 0.04;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
||| ||| :|:|:|:  
Db 592 LLGIWCGSKLICTTA 607

Search completed: May 7, 2004, 17:43:56  
Job time : 5.03361 sec

THE END

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 17:33:25 ; Search time 20.3025 Seconds  
(without alignments)  
248.653 Million cell updates/sec

Title: US-09-147-362A-9

Perfect score: 91  
Sequence: 1 LLSSWGCKGRLVCYTS 16

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	96.7	116	15 Q7ZJN9	Q7zjn9 human immun
2	88	96.7	216	15 Q9IEC5	Q9iec5 human immun
3	88	96.7	219	15 Q9IEB6	Q9ieb6 human immun
4	88	96.7	890	15 Q8Q7G2	Q8q7g2 human immun
5	87	95.6	218	15 Q9IE95	Q9ie95 human immun
6	87	95.6	538	15 Q9IED5	Q9ied5 human immun
7	86	94.5	219	15 Q9IEC8	Q9iec8 human immun
8	85	93.4	130	15 Q9IHU9	Q9ihu9 human immun
9	85	93.4	216	15 Q9IEC7	Q9iec7 human immun
10	85	93.4	535	15 Q9IEF2	Q9ief2 human immun
11	85	93.4	872	15 Q8Q7H0	Q8q7h0 human immun
12	85	93.4	882	15 Q8Q7F9	Q8q7f9 human immun
13	85	93.4	887	15 Q8Q7H6	Q8q7h6 human immun
14	85	93.4	887	15 Q8Q7G9	Q8q7g9 human immun
15	84	92.3	135	15 Q9DQL9	Q9dql9 human immun
16	84	92.3	209	15 Q9IE66	Q9ie66 human immun

17	84	92.3	242	15 Q9IE31	Q9ie31 human immun
18	82	90.1	115	15 Q7ZJN8	Q7zjn8 human immun
19	82	90.1	116	15 Q40459	Q40459 human immun
20	82	90.1	118	15 Q40451	Q40451 human immun
21	82	90.1	120	15 Q9IHU2	Q9ihu2 human immun
22	82	90.1	131	15 Q9WR05	Q9wr05 human immun
23	82	90.1	131	15 Q9IHU0	Q9ihu0 human immun
24	82	90.1	137	15 Q9IHV5	Q9ihv5 human immun
25	82	90.1	146	15 Q9WRV2	Q9wrv2 human immun
26	82	90.1	153	15 Q8J3R1	Q8j3r1 human immun
27	82	90.1	213	15 Q9IEC3	Q9iec3 human immun
28	82	90.1	214	15 Q9DIK3	Q9dik3 human immun
29	82	90.1	214	15 Q9IE96	Q9ie96 human immun
30	82	90.1	216	15 Q9IEA5	Q9iea5 human immun
31	82	90.1	230	15 Q9IEA9	Q9iea9 human immun
32	82	90.1	234	15 Q9IEA6	Q9iea6 human immun
33	82	90.1	342	15 Q11942	Q11942 human immun
34	82	90.1	529	15 Q9IEE2	Q9iee2 human immun
35	82	90.1	532	15 Q9IEE9	Q9iee9 human immun
36	82	90.1	544	15 Q9IED9	Q9ied9 human immun
37	82	90.1	548	15 Q9IED6	Q9ied6 human immun
38	82	90.1	551	15 Q9IEE1	Q9iee1 human immun
39	82	90.1	863	15 Q77377	Q77377 human immun
40	82	90.1	867	15 Q8Q7G8	Q8q7g8 human immun
41	82	90.1	871	15 Q57073	Q57073 human immun
42	82	90.1	871	15 Q57074	Q57074 human immun
43	82	90.1	871	15 Q8Q7I2	Q8q7i2 human immun
44	82	90.1	872	15 Q900Y5	Q900y5 human immun
45	82	90.1	872	15 Q57072	Q57072 human immun

ALIGNMENTS

RESULT 1

Q7ZJN9 PRELIMINARY; PRT; 116 AA.  
ID Q7ZJN9;  
AC Q7ZJN9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]\_TaxID=11676;  
RP SEQUENCE FROM N.A.  
RC STRAIN=08692A;  
RA Swanson P.A.; Devare S.G.; Hackett J.R. Jr.;  
RT "Molecular Characterization of 39 HIV-1 Isolates Representing Group M  
(Subtypes A-G) and Group O: Sequence Analysis of gag p24, pol  
integrase, and env gp41.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY214120; AAO61840.1; -  
DR GO; GO:0019031; C:viral envelope; IEA.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 13865 MW; AFIDFC578059061 CRC64;

Query Match 96.7%; Score 88; DB 15; Length 116;  
Best Local Similarity 93.8%; Pred. No. 3.5e-07;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
||:|||||  
Db 32 LLNSWGCKGRLVCYTS 47

RESULT 2

Q9IEC5 PRELIMINARY; PRT; 216 AA.  
ID Q9IEC5

AC Q9IEC5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp41 (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF07;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236394; CAB96243.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 216 216  
 SQ SEQUENCE 216 AA; 25027 MW; 413AE9BBF1B4FC9A CRC64;

Query Match 96.7%; Score 88; DB 15; Length 216;  
 Best Local Similarity 93.8%; Pred. No. 6.5e-07;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
 ||:|||||||  
 Db 47 LLNSWGCKGRLVCYIS 62

RESULT 3  
 Q9IEB6 PRELIMINARY; PRT; 219 AA.  
 AC Q9IEB6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp41 (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF14;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236404; CAB96252.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 219 219  
 SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 96.7%; Score 88; DB 15; Length 219;  
 Best Local Similarity 93.8%; Pred. No. 6.6e-07;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
 ||:|||||||  
 Db 57 LLNSWGCKGRLVCYTS 72

RESULT 4  
 Q8Q7G2 PRELIMINARY; PRT; 890 AA.  
 AC Q8Q7G2;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Envelope glycoprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=97US08692A;  
 RX MEDLINE=21849375; PubMed=11860674;  
 RA Yanaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
 RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
 RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
 RT Phylogenetic Clusters."  
 RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
 DR EMBL; AF383259; AAL98881.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 890 AA; 100610 MW; BBE816BC5C45EE23 CRC64;

Query Match 96.7%; Score 88; DB 15; Length 890;  
 Best Local Similarity 93.8%; Pred. No. 2.7e-06;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
 ||:|||||||  
 Db 620 LLNSWGCKGRLVCYTS 635

RESULT 5  
 Q9IE95 PRELIMINARY; PRT; 218 AA.  
 AC Q9IE95;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp41 (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YBF51;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236425; CAB96273.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 218 218  
 SQ SEQUENCE 218 AA; 25243 MW; E7F0E1E20EF79FA8 CRC64;

```
Query Match          95.6%; Score 87; DB 15; Length 218;
Best Local Similarity 87.5%; Pred. No. 9.8e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LLSSWGCKGRLVCYTS 16
        |||||:|||||:|||||
Db      54 LLSSWGCKGRLVCYTS 69

RESULT 6
Q9IEC5
ID      Q9IEC5      PRELIMINARY;      PRT;      538 AA.
AC      Q9IEC5;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Env polypeptide (Fragment).
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=YBF51;
RA      Roques P., Robertson D., Damond F., Sousquiere S., Mauciere P.,
RA      Depienne C., Brun-Vezinet F., Dormont D.;
RT      "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL      Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL, AJ133072; CAB96233.1; -.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP41; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT      NON_TER      1
FT      NON_TER      538
SQ      SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEF1D CRC64;

Query Match          95.6%; Score 87; DB 15; Length 538;
Best Local Similarity 87.5%; Pred. No. 2.4e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 LLSSWGCKGRLVCYTS 16
        |||||:|||||:|||||
Db      374 LLSSWGCKGRLVCYTS 389

RESULT 7
Q9IEC8
ID      Q9IEC8      PRELIMINARY;      PRT;      219 AA.
AC      Q9IEC8;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Gp41 (Fragment).
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BCF02;
RA      Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA      Philippe M.;
RT      "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL      Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL, AJ236391; CAB96240.1; -.
DR      GO; GO:0016021; C:integral to membrane; IEA.
```

```
DR      GO; GO:0019031; F:structural molecule activity; IEA.
DR      InterPro; IPR000328; Env GP41.
DR      Pfam; PF00517; GP41; 1.
KW      Transmembrane.
FT      NON_TER      1
FT      NON_TER      219
SQ      SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;

Query Match          94.5%; Score 86; DB 15; Length 219;
Best Local Similarity 87.5%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 LLSSWGCKGRLVCYTS 16
        |||||:|||||:|||||
Db      56 LLSSWGCKGRLVCYTS 71

RESULT 8
Q9IHU9
ID      Q9IHU9      PRELIMINARY;      PRT;      130 AA.
AC      Q9IHU9;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Envelope glycoprotein (Fragment).
GN      GP41.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=97CM798;
RX      MEDLINE=20386754; PubMed=10933623;
RA      Yang C., Gao F., Fonjongo P.N., Zekeng L., van der Groen G.,
RA      Fieniazek D., Schable C., Lal R.B.;
RT      "Phylogenetic analysis of protease and transmembrane regions of HIV
RT      type 1 group O.";
RL      AIDS Res. Hum. Retroviruses 16:1075-1081(2000).
DR      EMBL, AF229235; AAF71912.1; -.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019031; F:structural molecule activity; IEA.
DR      InterPro; IPR000328; Env GP41.
DR      Pfam; PF00517; GP41; 1.
KW      Transmembrane.
FT      NON_TER      1
FT      NON_TER      130
SQ      SEQUENCE 130 AA; 15593 MW; 5385789A336344EA CRC64;

Query Match          93.4%; Score 85; DB 15; Length 130;
Best Local Similarity 87.5%; Pred. No. 1.3e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 LLSSWGCKGRLVCYTS 16
        |||||:|||||:|||||
Db      40 LLSSWGCKGRLVCYTS 55

RESULT 9
Q9IEC7
ID      Q9IEC7      PRELIMINARY;      PRT;      216 AA.
AC      Q9IEC7;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Gp41 (Fragment).
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
```

```
RC STRAIN=BCF03;
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,
RA Philippe M.;
RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236392; CAB96241.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 25015 MW; 780C5F39CA6997C2 CRC64;

Query Match 93.4%; Score 85; DB 15; Length 216;
Best Local Similarity 93.8%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCYTS 16
   ||| ||||| ||||| |||||
DB 55 LLSWGCKGRVLCYTS 70

RESULT 10
ID Q9IEF2 PRELIMINARY; PRT; 535 AA.
AC Q9IEF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polypeptide (fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RA Roques P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
RA Deplenne C., Brun-Vezinet F., Dormont D.;
RT "HIV-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133055; CAB96216.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 535
SQ SEQUENCE 535 AA; 59682 MW; 7234BB8378DD12C5 CRC64;

Query Match 93.4%; Score 85; DB 15; Length 535;
Best Local Similarity 93.8%; Pred. No. 5.2e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCYTS 16
   ||| ||||| ||||| |||||
DB 374 LLSWGCKGRVLCYTS 389

RESULT 11
ID Q8Q7H0 PRELIMINARY; PRT; 872 AA.
AC Q8Q7H0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA453;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383251; AAL98873.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 872 AA; 98344 MW; 91D37F60053827E3 CRC64;

Query Match 93.4%; Score 85; DB 15; Length 872;
Best Local Similarity 87.5%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSWGCKGRVLCYTS 16
   ||| ||||| ||||| |||||
DB 602 LLSWGCKGRVLCYTS 617

RESULT 12
ID Q8Q7F9 PRELIMINARY; PRT; 882 AA.
AC Q8Q7F9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=98CMA100;
RX MEDLINE=21849375; PubMed=11860674;
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five
RT Phylogenetic Clusters.";
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).
DR EMBL; AF383262; AAL98884.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 882 AA; 99439 MW; 627BCA5D0C09C6D4 CRC64;

Query Match 93.4%; Score 85; DB 15; Length 882;
Best Local Similarity 87.5%; Pred. No. 8.5e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```



QY 1 LLSSWGCKGRLVCYTS 16  
DB 613 LLNSWGCKGRLVCYTA 628

## RESULT 13

Q8Q7H6 PRELIMINARY; PRT; 887 AA.  
AC Q8Q7H6; (1) :|||||:|||||:  
DT 01-JUN-2002 (TtEMBLrel. 21, Created)  
DT 01-JUN-2002 (TtEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TtEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98CWA124;  
RX MEDLINE=21849375; PubMed=11860674;  
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
RT Phylogenetic Clusters."  
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
DR EMBL; AF383245; AAL98867.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 887 AA; 9366 MW; E210F1E3F7B2474D CRC64;

Query Match 93.4%; Score 85; DB 15; Length 887;  
Best Local Similarity 87.5%; Pred. No. 8.5e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
DB 618 LLNSWGCKGRLVCYTS 633

## RESULT 14

Q8Q7G9 PRELIMINARY; PRT; 887 AA.  
AC Q8Q7G9; (1) :|||||:|||||:  
DT 01-JUN-2002 (TtEMBLrel. 21, Created)  
DT 01-JUN-2002 (TtEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TtEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=98CWA124;  
RX MEDLINE=21849375; PubMed=11860674;  
RA Yamaguchi J., Vallari A.S., Swanson P., Bodelle P., Kaptue L.,  
RA Ngansop C., Zekeng L., Gurtler L.G., Devare S.G., Brennan C.A.;  
RT "Evaluation of HIV Type 1 Group O Isolates: Identification of Five  
RT Phylogenetic Clusters."  
RL AIDS Res. Hum. Retroviruses 18:269-282(2002).  
DR EMBL; AF383252; AAL98874.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 887 AA; 9559 MW; 88954F4ED76A314A CRC64;

Query Match 93.4%; Score 85; DB 15; Length 887;  
Best Local Similarity 87.5%; Pred. No. 8.5e-06;  
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
DB 618 LLNSWGCKGRLVCYTS 633

## RESULT 15

Q9DQL9 PRELIMINARY; PRT; 135 AA.  
AC Q9DQL9;  
DT 01-MAR-2001 (TtEMBLrel. 16, Created)  
DT 01-MAR-2001 (TtEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TtEMBLrel. 24, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP-M1321;  
RX MEDLINE=20584646; PubMed=11153079;  
RA Ortiz M., Munoz L., Bernal A., Rodriguez A., Zorraquino A.,  
RA Vellido J., Salas A., Moreno A., Garcia-Saiz A.;  
RT "Molecular Characterization of Non-B HIV Type 1 Subtypes from Africa  
RT in Spain."  
RL AIDS Res. Hum. Retroviruses 16:1967-1971(2000).  
DR EMBL; AF255939; AAG36894.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
DR Pfam; PF00516; GP120; 1.  
FT NON\_TER 1 135  
FT NON\_TER 135 135  
SQ SEQUENCE 135 AA; 16285 MW; D34B9EF77318BB36 CRC64;

Query Match 92.3%; Score 84; DB 15; Length 135;  
Best Local Similarity 81.2%; Pred. No. 2e-06;  
Matches 13; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSSWGCKGRLVCYTS 16  
DB 48 LLNSWGCKGRLVCYTS 63

Search completed: May 7, 2004, 17:49:13  
Job time : 21.3025 secs

THE UNIVERSITY OF CHICAGO